



Snustad | Simmons

Fundamentos de
GENÉTICA

6ª edição




GUANABARA
KOOGAN

Fundamentos de Genética | 6ª edição integra teoria e prática, e apresenta um equilíbrio entre novas informações e conceitos fundamentais.

O texto contém recursos especiais destinados a enfatizar a relevância dos tópicos expostos, facilitar a compreensão de conceitos importantes e ajudar os estudantes a avaliarem seus conhecimentos.

Características da obra

- Boxes intitulados **Problema resolvido** apresentam estratégias para abordar e resolver questões essenciais
- Cada capítulo conta com três seções de exercícios – **Exercício**, **Autoavaliação** e **Avaliação adicional** –, para que o leitor aplique a teoria estudada
- Boxes **Em foco** abrangem os mais recentes avanços da Genética e seus tópicos mais relevantes
- Boxes **Resolva!** apresentam questões que estimulam a pesquisa dos temas abordados.

Material suplementar

- Explicações, passo a passo, das respostas dos problemas do boxe **Resolva!**
- Todas as respostas da seção **Autoavaliação** (acesso restrito a docentes).



www.grupogen.com.br
<http://gen-io.grupogen.com.br>

ISBN 978-85-277-2277-3



9 788527 722773



0009182659

Fundamentos de Genética


■ Os autores deste livro e a EDITORA GUANABARA KOOGAN LTDA. empenharam seus melhores esforços para assegurar que as informações e os procedimentos apresentados no texto estejam em acordo com os padrões aceitos à época da publicação, *e todos os dados foram atualizados pelos autores até a data da entrega dos originais à editora*. Entretanto, tendo em conta a evolução das ciências da saúde, as mudanças regulamentares governamentais e o constante fluxo de novas informações sobre terapêutica medicamentosa e reações adversas a fármacos, recomendamos enfaticamente que os leitores consultem sempre outras fontes fidedignas, de modo a se certificarem de que as informações contidas neste livro estão corretas e de que não houve alterações nas dosagens recomendadas ou na legislação regulamentadora. *Adicionalmente, os leitores podem buscar por possíveis atualizações da obra em <http://gen-io.grupogen.com.br>.*

■ Os autores e a editora se empenharam para citar adequadamente e dar o devido crédito a todos os detentores de direitos autorais de qualquer material utilizado neste livro, dispondo-se a possíveis acertos posteriores caso, inadvertida e involuntariamente, a identificação de algum deles tenha sido omitida.

■ Traduzido de:
PRINCIPLES OF GENETICS, SIXTH EDITION
Copyright © 2012, 2009, 2006, 2003, 2000, and 1997 John Wiley & Sons, Inc.
All Rights Reserved. This translation published under license.
ISBN: 978-1-11812921-0

■ Direitos exclusivos para a língua portuguesa
Copyright © 2013 by
EDITORA GUANABARA KOOGAN LTDA.
Uma editora integrante do GEN | Grupo Editorial Nacional
Travessa do Ouvidor, 11
Rio de Janeiro – RJ – CEP 20040-040
Tels.: (21) 3543-0770/(11) 5080-0770 | Fax: (21) 3543-0896
www.editoraguanabara.com.br | www.grupogen.com.br | editorial.saude@grupogen.com.br

■ Reservados todos os direitos. É proibida a duplicação ou reprodução deste volume, no todo ou em parte, em quaisquer formas ou por quaisquer meios (eletrônico, mecânico, gravação, fotocópia, distribuição pela Internet ou outros), sem permissão, por escrito, da EDITORA GUANABARA KOOGAN LTDA.

■ Capa: Renato de Mello
Editoração eletrônica:  Anthares

■ Ficha catalográfica

S651f
6.ed.

Snustad, D. Peter, 1940-
Fundamentos de genética / D. Peter Snustad, Michael J. Simmons; tradução Cláudia Lúcia Caetano de Araújo. – 6. ed. – Rio de Janeiro: Guanabara Koogan, 2013.

Tradução de: Principles of genetics
ISBN 978-85-277-2277-3

1. Genética humana. I. Simmons, Michael J. II. Título.

13-1493.

CDD: 576.5
CDU: 575

Dedicatória

A Judy, minha mulher e melhor amiga.
D.P.S.

A minha família, especialmente a Benjamin.
M.J.S.

Sobre os Autores

D. Peter Snustad é Professor Emérito na University of Minnesota, Twin Cities. Graduado pela University of Minnesota e pós-graduado (mestrado e doutorado) pela University of California, Davis. Iniciou a carreira de docente no Department of Agronomy and Plant Genetics, em Minnesota em 1965, tornou-se membro fundador do novo Department of Genetics em 1966 e transferiu-se para o Department of Plant Biology em 2000. Durante seus 43 anos em Minnesota, ministrou cursos sobre vários temas, desde a biologia geral até a genética bioquímica. Sua pesquisa inicial concentrou-se nas interações entre o bacteriófago T4 e seu hospedeiro, a *E. coli*. Na década de 1980, passou a pesquisar o citoesqueleto de *Arabidopsis* e os genes da glutamina sintetase do milho. Dentre outras homenagens, recebeu os prêmios Morse-Amoco e Dagley Memorial para o ensino, e foi eleito para *Fellow* da American Association for the Advancement of Science. O amor pela natureza canadense manteve-o próximo à Minnesota.

Michael J. Simmons é Professor no Department of Genetics, Cell Biology and Development na University of Minnesota, Twin Cities. Graduou-se em Biologia em St. Vincent College, em Latrobe, Pennsylvania, e concluiu o mestrado e o doutorado em Genética na University of Wisconsin, Madison. Dr. Simmons já ministrou vários cursos, entre eles Genética e Genética de Populações. Também orientou muitos estudantes em projetos de pesquisa. No início da carreira recebeu o prêmio Morse-Amoco para o ensino da University of Minnesota em reconhecimento a suas contribuições para o ensino universitário. A pesquisa do Dr. Simmons concentra-se no significado genético de elementos transponíveis no genoma de *Drosophila melanogaster*. Ele atuou em comitês consultivos no National Institutes of Health e foi membro do corpo editorial da revista *Genetics* durante 21 anos. Uma de suas atividades favoritas, a patinação artística, é particularmente compatível com o clima de Minnesota.

Material Suplementar

Este livro conta com o seguinte material suplementar:

- Explicações passo a passo das respostas dos problemas do box *Resolva!*
- Todas as respostas da seção *Autoavaliação* (restrito a docentes)

Para ter acesso a esse conteúdo, que é gratuito, o docente ou leitor deve se cadastrar em <http://gen-io.grupogen.com.br>. Além disso, para que este material específico seja válido, é necessário que se informe o código existente na etiqueta colada na parte interna da capa do livro.



GEN-IO (GEN | Informação Online) é o repositório de materiais suplementares e de serviços relacionados com livros publicados pelo GEN | Grupo Editorial Nacional, maior conglomerado brasileiro de editoras do ramo científico-técnico-profissional, composto por Guanabara Koogan, Santos, Roca, AC Farmacêutica, Forense, Método, LTC, E.P.U. e Forense Universitária. Os materiais suplementares ficam disponíveis para acesso durante a vigência das edições atuais dos livros a que eles correspondem.

Prefácio

As pesquisas em genética vêm avançando rapidamente. É possível analisar com muitos detalhes o DNA dos genomas, mesmo os grandes, estudar as funções de genes individuais por meio de uma série de técnicas impressionantes, bem como modificar geneticamente organismos mediante introdução de genes estranhos ou alterados em seus genomas. Os métodos de ensino e aprendizado da genética também estão mudando. Muitos são os recursos eletrônicos para garantir acesso a informações e transmissoras, mídias novas e atraentes estão sendo desenvolvidas, e as salas de aula de diversas instituições estão sendo reformadas para incorporar estratégias de “aprendizado ativo”. Esta edição de *Fundamentos de Genética* foi elaborada para reconhecer esses avanços científicos e educacionais.

Objetivos

Fundamentos de Genética apresenta um equilíbrio entre novas informações e conceitos fundamentais. A preparação desta edição foi pautada em quatro objetivos principais:

- **Foco nos princípios básicos da genética**, com apresentação minuciosa e completa dos importantes conceitos da genética clássica, molecular e de populações. A base sólida é essencial para a compreensão dos avanços atuais da genética e o reconhecimento de seu significado prático. Além disso, a extensão e a profundidade da cobertura das diferentes áreas da genética – clássica, molecular e de populações – têm de ser equilibradas, e o volume de informações em permanente expansão na genética tem de ser organizado por um modelo forte, mas flexível, de conceitos-chave.
- **Foco no processo científico**, mostrando como se dá o desenvolvimento dos conceitos científicos a partir de observações e experimentos. Esta obra apresenta muitos exemplos de como os princípios genéticos emergiram do trabalho de diferentes cientistas. A ciência é enfatizada como um processo contínuo de observação, experimentação e descoberta.
- **Foco na genética humana**, incorporando exemplos humanos e mostrando a relevância da genética nas questões relativas à sociedade. A experiência mostra que os estudantes têm interesse especial pela genética da sua própria espécie e que, por isso, compreendem com mais facilidade os conceitos complexos quando ilustrados com exemplos humanos. Desse modo, sempre que possível, esses exemplos foram usados. Incluímos também discussões do Projeto Genoma Humano, mapeamento de genes humanos, distúrbios genéticos,

terapia gênica e aconselhamento genético em todo o texto. Questões como triagem genética, análise do perfil de DNA, engenharia genética, clonagem, pesquisa com células-tronco e terapia gênica incitaram densos debates sobre as ramificações sociais, jurídicas e éticas da genética. É importante que os estudantes sejam incluídos nas discussões sobre essas questões, e este livro garante-lhes os requisitos para tal.

- **Foco no desenvolvimento de habilidades de reflexão**, enfatizando a análise de dados experimentais e problemas. A genética sempre foi um pouco diferente de outras disciplinas na Biologia em vista da forte ênfase na solução de problemas. Neste texto, a natureza analítica da genética é detalhada de muitas maneiras – o desenvolvimento de princípios na genética clássica, a discussão dos experimentos na genética molecular e a apresentação dos cálculos na genética de populações. Em todo o livro, enfatiza-se a integração dos dados de observação e experimentais com a análise lógica para o desenvolvimento de conceitos-chave. Cada capítulo tem dois grupos de problemas resolvidos – a seção *Exercícios*, que contém problemas simples de aplicação da análise genética básica, e a seção *Autoavaliação*, que contém problemas mais complexos que integram diferentes conceitos e técnicas. Um conjunto de *Avaliações adicionais* acompanha os problemas resolvidos, de modo que os estudantes compreendam melhor os conceitos do capítulo e desenvolvam a capacidade de análise. Outra seção, *Genômica na Web*, apresenta questões que podem ser respondidas por meio de pesquisa no *site* do National Center for Biotechnology Information. Nesta seção, os estudantes aprendem a usar o grande repositório de informações genéticas acessíveis no *site* e podem aplicar as informações em problemas específicos. Cada capítulo conta ainda com uma seção, *Problema resolvido*, que propõe um problema, lista os fatos e conceitos pertinentes, analisa-o e apresenta sua solução. Por fim, outra novidade: os boxes *Resolva!* garantem aos estudantes a oportunidade de testar a compreensão dos conceitos estudados. As respostas para as questões apresentadas nos boxes *Resolva!* estão disponíveis no *site* <http://gen-io.grupogen.com.br>.

Conteúdo e organização da sexta edição

A organização desta edição de *Fundamentos de Genética* é semelhante à anterior. O conteúdo, porém, foi depu-

rado e atualizado para tornar possível a atualização cuidadosa. Ao selecionarmos o material a ser incluído nesta edição, tentamos ser abrangentes, mas não enciclopédicos.

O texto foi dividido em 24 capítulos, um a menos que a edição anterior. Os Capítulos 1 e 2 apresentam a ciência da genética, aspectos básicos da reprodução celular e alguns organismos genéticos-modelo; os Capítulos 3 a 8 apresentam os conceitos da genética clássica e os procedimentos básicos da análise genética de microrganismos; os Capítulos 9 a 13 apresentam os tópicos da genética molecular, entre eles replicação do DNA, transcrição, tradução e mutação; os Capítulos 14 a 17 abordam tópicos mais avançados de genética molecular e genômica; os Capítulos 18 a 21 tratam da regulação da expressão gênica e da base genética do desenvolvimento, da imunidade e do câncer; os Capítulos 22 a 24 apresentam os conceitos de genética quantitativa, de populações e evolutiva.

Como nas edições anteriores, tentamos criar um texto adaptável a diferentes formatos de curso. Muitos professores preferem apresentar os tópicos da mesma maneira que apresentamos, começando com a genética clássica, avançando para a genética molecular e terminando com a genética quantitativa, de populações e evolutiva. No entanto, o texto foi elaborado de modo que os docentes possam apresentar os tópicos em diferentes ordens. Eles podem, por exemplo, começar com genética molecular básica (Capítulos 9 a 13), depois apresentar a genética clássica (Capítulos 3 a 8), passar a tópicos mais avançados de genética molecular (Capítulos 14 a 21) e terminar o curso com a genética quantitativa, de populações e evolutiva (Capítulos 22 a 24). Outra opção é inserir a genética quantitativa e de populações entre a genética clássica e a molecular.

Pedagogia da sexta edição

O texto contém recursos especiais destinados a enfatizar a relevância dos tópicos expostos, facilitar a compreensão de conceitos importantes e ajudar os estudantes a avaliarem seus conhecimentos.

- **Narrativa de abertura do capítulo.** Cada capítulo inicia-se com um breve texto que destaca o significado dos tópicos apresentados.
- **Sumário do capítulo.** As principais seções são apresentadas, de maneira conveniente, na primeira página de cada capítulo.
- **Resumo da seção.** Há um breve resumo do conteúdo no início de cada seção principal do texto. Esses resumos introdutórios concentram a atenção nas principais ideias expostas no capítulo.
- **Pontos essenciais.** Esses recursos estão no fim de cada seção principal de um capítulo. O objetivo é ajudar os alunos a estudarem para as provas e a recapitularem as principais ideias do conteúdo estudado.
- **Em foco.** Os boxes *Em foco* apresentam tópicos especiais. O conteúdo desses boxes reforça ou expande conceitos, técnicas ou habilidades descritos no texto.
- **O futuro.** O conteúdo desses boxes destaca novas e empolgantes descobertas em genética – com frequência, tema de pesquisas correntes.
- **Problema resolvido.** Cada capítulo conta com um box que orienta o estudante na análise e na solução de um problema que inclua conteúdo importante do capítulo. O box lista fatos e conceitos relevantes para o problema e, depois, explica como resolvê-lo.
- **Resolva!** Cada capítulo contém dois boxes *Resolva!* Neste box, há sempre um problema relacionado com os conceitos apresentados no texto, possibilitando a avaliação da compreensão de conceitos-chave. A solução, passo a passo, de cada problema está disponível no site <http://gen-io.grupogen.com.br>.
- **Exercícios.** No fim de cada capítulo, apresentamos vários problemas resolvidos para reforçar os conceitos fundamentais apresentados. O objetivo desses exercícios simples, em uma etapa, é ilustrar a análise genética básica ou destacar informações importantes.
- **Autoavaliação.** Cada capítulo também tem problemas resolvidos mais complexos para ajudar os estudantes a aprimorarem a capacidade de análise e solução de problemas. Os problemas dessa seção destinam-se a integrar diferentes conceitos e técnicas. Na análise de cada problema, mostramos o passo a passo da solução.
- **Avaliação adicional.** Cada capítulo termina com um conjunto de questões e problemas de dificuldades variadas, organizados de acordo com a sequência de tópicos no capítulo. As questões e os problemas mais difíceis são identificados por números coloridos. Esses conjuntos de questões e problemas dão aos estudantes a oportunidade de compreender melhor os conceitos abordados no capítulo e desenvolver a capacidade de análise.
- **Genômica na web.** As informações sobre genomas, genes, sequências de DNA, organismos mutantes, sequências de polipeptídeos, vias bioquímicas e relações evolutivas estão disponíveis gratuitamente em diversos sites na internet. A consulta a essas informações faz parte da rotina dos pesquisadores, e nós acreditamos que os estudantes devem se familiarizar com elas. Com esse objetivo, ao fim de cada capítulo, foi incluída uma série de questões que podem ser respondidas por consulta ao site do National Center for Biotechnology Information (NCBI), patrocinado pelo National Institutes of Health dos Estados Unidos.
- **Apêndices.** Cada Apêndice apresenta conteúdo técnico útil para análise genética.
- **Glossário.** Nesta seção são definidos termos importantes. É um ótimo recurso para esclarecer alguns tópicos e se preparar para provas.
- **Respostas.** As respostas aos itens de número ímpar da seção *Avaliação adicional* podem ser encontradas no fim do livro.

Agradecimentos

Como as anteriores, esta edição de *Fundamentos de Genética* teve muita influência dos cursos em que lecionamos. Agradecemos a nossos alunos, pelas opiniões construtivas sobre o conteúdo e o método pedagógico, e a nossos colegas da University of Minnesota, por compartilharem seu conhecimento e experiência. Professores de Genética de outras instituições também contribuíram com muitas sugestões úteis. Agradecemos especialmente aos revisores desta e das edições anteriores, destacados a seguir.

Revisores da 6ª edição

Ann Aguano, Manhattan Marymount College; Mary A. Bedell, University of Georgia; Jonathan Clark, Weber State University; Robert Fowler, San Jose State University; Cheryl Hertz, Loyola Marymount University; Shawn Kaeppler, University of Wisconsin; Todd Kelson, Brigham Young University – Idaho; Richard D. Noyes, University of Central Arkansas; Maria E. Orive, University of Kansas; Rongsun Pu, Kean University.

Revisores das edições anteriores

Michelle Boissere, Xavier University of Louisiana; Stephen P. Bush, Coastal Carolina University; Sarah Crawford, Southern Connecticut State University; Xiongbin Lu, University of South Carolina – Columbia; Valery N. Soyfer, George Mason University; David Starkey, University of Central Arkansas; Frans Tax, University of Arizona; Tzvi Tzfira, University of Michigan; Harald Vaessin, The

Ohio State University – Columbus; Sarah VanVickle-Chavez, Washington University in St. Louis; Willem Vermeris, University of Florida; Alan S. Waldman, University of South Carolina – Columbia.

Muitas pessoas contribuíram para a elaboração e produção desta edição. Kevin Witt, Editor Sênior, e Michael Palumbo, Editor Assistente, iniciaram o projeto e deram ideias para alguns aspectos do texto. A Dra. Pamela Marshall, da Arizona State University, sugeriu muitas maneiras de aperfeiçoar a edição anterior, e um grupo de professores de genética comentou cuidadosamente suas sugestões. Os membros desse grupo foram: Anna Aguano, Manhattan Marymount College; Robert Fowler, San Jose State University; Jane Glazebrook, University of Minnesota; Shawn Kaeppler, University of Wisconsin; Todd Kelson, Brigham Young University – Idaho; e Dwayne A. Wise, Mississippi State University. Agradecemos a esses experientes professores de genética pela contribuição.

Jennifer Dearden e Lauren Morris ajudaram em muitos detalhes logísticos na preparação desta edição, e Lisa Passmore pesquisou e conseguiu muitas fotografias novas. Jennifer MacMillan, Editora de Fotografia Sênior, coordenou com competência todo o programa de fotos. Somos gratos pela contribuição de todos. Agradecemos a Maureen Eide, Designer Sênior, por criar um novo *layout* do texto, e a Precision Graphics e Aptara por fazerem as ilustrações. Elizabeth Swain, Editora de Produção Sênior, coordenou fantasticamente a produção desta edição; Betty Pessagno preparou fielmente os originais; Lilian Brady fez a leitura de prova e Stephen Ingle preparou o índice alfabético. Somos muito gratos a todas essas pessoas pelo excelente trabalho. Agradecemos também a Clay Stone, Gerente Executivo de *Marketing*, que envidou seus esforços para que esta edição chegasse às mãos dos leitores.

Sumário

Capítulo 1

Ciência da Genética, 1

Um convite, 2

Três grandes marcos da genética, 2

Mendel | Os genes e as regras da herança, 2

Watson e Crick | A estrutura do DNA, 3

Projeto Genoma Humano | Sequenciamento do DNA e catalogação dos genes, 4

DNA como material genético, 6

Replicação do DNA | Propagação da informação genética, 6

Expressão gênica | Uso da informação genética, 6

Mutação | Mudança na informação genética, 8

Genética e evolução, 10

Níveis de análise genética, 10

Genética clássica, 11

Genética molecular, 11

Genética de populações, 11

Genética no mundo | Aplicações da genética nos empreendimentos humanos, 12

Genética na agricultura, 12

Genética na medicina, 13

Genética na sociedade, 14

Capítulo 2

Reprodução Celular, 17

Células e cromossomos, 18

Ambiente celular, 18

Células procarióticas e eucarióticas, 18

Cromossomos | Onde estão localizados os genes, 20

Divisão celular, 21

Mitose, 22

Meiose, 25

Meiose I, 26

Meiose II e os resultados da meiose, 30

Ciclos de vida de alguns organismos genéticos-modelo, 31

Saccharomyces cerevisiae, fermento para pão, 31

Arabidopsis thaliana, uma planta de crescimento rápido, 32

Mus musculus, o camundongo, 32

Capítulo 3

Mendelismo | Princípios Básicos da Herança, 38

Estudos de Mendel sobre a hereditariedade, 39

Organismo experimental de Mendel, a ervilha, 39

Cruzamentos mono-híbridos | Os princípios da dominância e da segregação, 39

Cruzamentos di-híbridos | O princípio da distribuição independente, 42

Aplicações dos princípios de Mendel, 44

Método do quadrado de Punnett, 44

Método da linha bifurcada, 44

Método da probabilidade, 44

Teste das hipóteses genéticas, 47

Teste do qui-quadrado, 47

Princípios mendelianos em genética humana, 50

Heredogramas, 50

Segregação mendeliana em famílias humanas, 51

Aconselhamento genético, 52

Capítulo 4

Extensões do Mendelismo, 61

Variação alélica e função gênica, 62

Dominância incompleta e codominância, 62

Alelos múltiplos, 63

Série alélica, 63

Testando mutações gênicas para alelismo, 64

Variação entre os efeitos das mutações, 65

Funções dos genes na produção de polipeptídios, 66

Por que algumas mutações são dominantes e outras recessivas?, 67

Ação gênica | Do genótipo ao fenótipo, 69

Influência do ambiente, 69

Efeitos ambientais sobre a expressão de genes humanos, 70

Penetrância e expressividade, 70

Interações gênicas, 70

Epistasia, 71

Pleiotropia, 74

Endogamia | Outro olhar nos heredogramas, 76

Efeitos da endogamia, 76

Análise genética da endogamia, 77

Medida das relações genéticas, 79

Capítulo 5

Base Cromossômica do Mendelismo, 88

Cromossomos, 89

Número de cromossomos, 89

Cromossomos sexuais, 89

Teoria cromossômica da hereditariedade, 90

Evidências experimentais que associam

a herança de genes aos cromossomos, 91

Não disjunção como comprovação da teoria cromossômica, 91

Base cromossômica dos princípios de segregação e distribuição independente de Mendel, 93

Genes ligados ao sexo em seres humanos, 96

Hemofilia, distúrbio da coagulação sanguínea ligado ao X, 97

Discromatopsia, um distúrbio da visão ligado ao X, 98

Genes no cromossomo Y humano, 98

Genes nos cromossomos X e Y, 99

Cromossomos sexuais e determinação do sexo, 99

Determinação do sexo em seres humanos, 99

Determinação do sexo em *Drosophila*, 100

Determinação do sexo em outros animais, 100

Compensação de dose de genes ligados ao X, 102

Hiperativação de genes ligados ao X em machos de *Drosophila*, 102

Inativação de genes ligados ao X em fêmeas de mamíferos, 102

Capítulo 6

Variação no Número e na Estrutura dos Cromossomos, 109

Técnicas citológicas, 110

Análise de cromossomos mitóticos, 110

Cariótipo humano, 111

Variação citogenética | Considerações gerais, 112

Poliploidia, 113

Poliploides estéreis, 114

Poliploides férteis, 114

Poliploidia e politenia tecido-específica, 116

Aneuploidia, 118

Trissomia em seres humanos, 118

Monossomia, 120

Deleções e duplicações de segmentos cromossômicos, 121

Rearranjos da estrutura do cromossomo, 124

Inversões, 125

Translocações, 125

Cromossomos compostos e translocações robertsonianas, 127

Capítulo 7

Ligação, Crossing over e Mapeamento Cromossômico em Eucariotos, 134

Ligação, recombinação e *crossing over*, 135

Evidências iniciais da ligação

e recombinação, 135

Crossing over como base física

da recombinação, 137

Evidências de que o *crossing over* causa recombinação, 139

Quiasmas e o momento do *crossing over*, 139

Mapeamento cromossômico, 140

Crossing over como medida de

distância genética, 140

Mapeamento de recombinação com um

cruzamento-teste de dois pontos, 141

Mapeamento de recombinação com um

cruzamento-teste de três pontos, 142

Frequência de recombinação e distância no mapa genético, 145

Mapeamento citogenético, 147

Localização de genes com o auxílio das deleções e duplicações, 147

Distância genética e distância física, 149

Análise de ligação em seres humanos, 150

Recombinação e evolução, 152

Significado evolutivo da recombinação, 152

Supressão da recombinação por inversões, 152

Controle genético da recombinação, 154

Capítulo 8

Genética de Bactérias e seus Vírus, 163

Vírus e bactérias em genética, 164

Genética dos vírus, 165

Bacteriófago T4, 165

Bacteriófago lambda, 166

Genética das bactérias, 169

Genes mutantes em bactérias, 170

Transferência gênica unidirecional em bactérias, 171

Mecanismos de troca genética em bactérias, 171

Transformação, 172

Conjugação, 174

Plasmídios e episômos, 178

Fatores F' e sexodução, 181

Transdução, 182

Significado evolutivo da troca genética em bactérias, 186

Capítulo 9

DNA e a Estrutura Molecular dos Cromossomos, 193

- Funções do material genético, 194
- Comprovação de que as informações genéticas são armazenadas no DNA, 194
 - Comprovação de que o DNA é o mediador da transformação, 194
 - Comprovação de que o DNA contém as informações genéticas no bacteriófago T2, 195
 - Comprovação de que o RNA armazena as informações genéticas em alguns vírus, 197
- Estruturas do DNA e do RNA, 198
 - Natureza das subunidades químicas no DNA e no RNA, 198
 - Estrutura do DNA | A dupla hélice, 198
 - Estrutura do DNA | Formas alternativas da dupla hélice, 204
 - Estrutura do DNA | Super-hélices negativas *in vivo*, 205
- Estrutura do cromossomo em procariotos e vírus, 206
- Estrutura do cromossomo em eucariotos, 207
 - Composição química de cromossomos eucarióticos, 208
 - Uma grande molécula de DNA por cromossomo, 209
 - Três níveis de empacotamento do DNA em cromossomos eucarióticos, 209
 - Centrômeros e telômeros, 212
 - Sequências repetidas de DNA, 215

Capítulo 10

Replicação do DNA e dos Cromossomos, 222

- Características básicas da replicação de DNA *in vivo*, 223
 - Replicação semiconservativa, 223
 - Origens únicas de replicação, 226
 - Visualização de forquilha de replicação por autorradiografia, 229
 - Replicação bidirecional, 230
- Replicação de DNA em procariotos, 232
 - Síntese contínua de um filamento e síntese descontínua do outro, 233
 - Fechamento covalente de cortes no DNA por DNA ligase, 234
 - Iniciação da replicação do DNA, 234
 - Iniciação de cadeias de DNA com iniciadores de RNA, 236
 - Desenrolamento de DNA com helicases, proteínas de ligação ao DNA e topoisomerases, 237

- Múltiplas DNA polimerases e revisão, 241
- O primossomo e o replissomo, 244
- Replicação por círculo rolante, 245

Aspectos específicos da replicação de cromossomos eucarióticos, 247

- Ciclo celular, 247
- Múltiplos réplicons por cromossomo, 247
- Duas ou mais DNA polimerases em uma única forquilha de replicação, 249
- Duplicação de nucleossomos nas forquilhas de replicação, 250
- Telomerase | Replicação das terminações do cromossomo, 251
- Comprimento do telômero e envelhecimento em seres humanos, 252

Capítulo 11

Transcrição e Processamento do RNA, 260

- Transferência de informações genéticas | O dogma central, 261
 - Transcrição e tradução, 261
 - Cinco tipos de moléculas de RNA, 263
- Processo de expressão gênica, 263
 - Um mRNA intermediário, 263
 - Aspectos gerais da síntese de RNA, 265
- Transcrição em procariotos, 267
 - RNA polimerases | Enzimas complexas, 267
 - Iniciação de cadeias de RNA, 268
 - Alongamento de cadeias de RNA, 268
 - Término das cadeias de RNA, 268
 - Transcrição, tradução e degradação de mRNA concomitantes, 269
- Transcrição e processamento de RNA em eucariotos, 271
 - Cinco RNA polimerases/cinco conjuntos de genes, 271
 - Iniciação de cadeias de RNA, 274
 - Alongamento da cadeia de RNA e acréscimo de *caps* de metilguanossina na extremidade 5', 276
 - Término por clivagem da cadeia e o acréscimo de caudas poli(A) 3', 277
 - Edição de RNA | Alteração das informações contidas nas moléculas de mRNA, 278
- Genes interrompidos em eucariotos | Éxons e íntrons, 279
 - Alguns genes eucarióticos muito grandes, 280
 - Íntrons | Significado biológico?, 281
- Remoção de sequências de íntrons por recomposição de RNA, 281
 - Recomposição do precursor de tRNA | Atividades específicas de nuclease e ligase, 282

- Recomposição autocatalítica, 282
- Recomposição de pré-mRNA |
 - snRNA, snRNP e spliceossomo, 284

Capítulo 12

Tradução e Código Genético, 290

- Estrutura das proteínas, 291
 - Polipeptídios | Vinte subunidades diferentes de aminoácidos, 291
 - Proteínas | Estruturas tridimensionais complexas, 291
- Gene | Um polipeptídio colinear, 294
 - Beadle e Tatum | Um gene – uma enzima, 294
 - Colinearidade entre a sequência codificadora de um gene e seu produto polipeptídico, 296
- Síntese proteica | Tradução, 297
 - Visão geral da síntese proteica, 297
 - Componentes necessários para a síntese proteica | Ribossomos, 298
 - Componentes necessários à síntese proteica | RNA transportador, 300
 - Tradução | A síntese de polipeptídios usando moldes de mRNA, 303
- Código genético, 309
 - Propriedades do código genético | Considerações gerais, 310
 - Três nucleotídeos por códon, 310
 - Decifrando o código, 312
 - Códons de iniciação e de término, 312
 - Um código degenerado e ordenado, 313
 - Um código quase universal, 315
- Interações códon-tRNA, 315
 - Reconhecimento de códons por tRNA | A hipótese da oscilação, 315
 - Mutações supressoras que produzem tRNA com reconhecimento do códon alterado, 316

Capítulo 13

Mutação, Reparo do DNA e Recombinação, 323

- Mutação | Origem da variabilidade genética necessária para a evolução, 324
- Base molecular da mutação, 324
 - Mutações induzidas, 325
 - Mutações induzidas por substâncias químicas, 328
 - Mutações induzidas por radiação, 332
 - Mutações induzidas por elementos genéticos transponíveis, 333
 - Expansão de repetições de trinucleotídeos e doenças humanas hereditárias, 333
- Mutação | Características básicas do processo, 335
 - Mutação | Somática ou germinativa, 335

- Mutação | Espontânea ou induzida, 336
- Mutação | Geralmente um processo aleatório, não adaptativo, 336
- Mutação | Um processo reversível, 338

Mutação | Efeitos fenotípicos, 339

- Mutações com efeitos fenotípicos |
 - Geralmente prejudiciais e recessivas, 339
- Efeitos das mutações nos genes da globina humana, 340
- Mutação em seres humanos | Bloqueios em vias metabólicas, 341
- Mutações letais condicionais | Instrumentos eficientes para estudos genéticos, 342

Localização das mutações nos genes pelo teste de complementação, 344

Avaliação da mutagenicidade de substâncias químicas | Teste de Ames, 347

Mecanismos de reparo do DNA, 349

- Reparo dependente de luz, 349
- Reparo por excisão, 349
- Outros mecanismos de reparo do DNA, 352

Doenças humanas hereditárias com defeitos no reparo do DNA, 353

Mecanismos de recombinação do DNA, 355

- Recombinação | Clivagem e reunião das moléculas de DNA, 355
- Conversão gênica | Síntese de reparo do DNA associada à recombinação, 358

Capítulo 14

Técnicas de Genética Molecular, 367

- Técnicas básicas para identificação, amplificação e clonagem de genes, 368
 - A descoberta das endonucleases de restrição, 369
 - Produção *in vitro* de moléculas de DNA recombinantes, 372
 - Amplificação de moléculas de DNA recombinantes em vetores de clonagem, 373
 - Clonagem de grandes genes e segmentos de genomas em BAC, PAC e YAC, 374
 - Amplificação das sequências de DNA pela reação em cadeia da polimerase (PCR), 375
- Construção e rastreamento das bibliotecas de DNA, 377
 - Construção de bibliotecas genômicas, 378
 - Construção de bibliotecas de cDNA, 379
 - Rastreamento de bibliotecas de DNA para identificar genes de interesse, 379
- Análise molecular de DNA, RNA e proteínas, 381
 - Análise de DNA por hibridizações *Southern blot*, 382
 - Análise de RNA por hibridizações *Northern blot*, 382
 - Análise de RNA por PCR com transcriptase reversa (RT-PCR), 382
 - Análise de proteínas por técnicas *Western blot*, 384

- Análise molecular de genes e cromossomos, 387
- Mapas físicos de moléculas de DNA baseados em sítios de clivagem por enzima de restrição, 387
- Sequências nucleotídicas de genes e cromossomos, 388

Capítulo 15

Genômica, 398

- Genômica | Considerações gerais, 402
- Correlação de mapas genéticos, citológicos e físicos de cromossomos, 403
- Mapas de polimorfismo do comprimento do fragmento de restrição (RFLP) e repetição curta em série (STR), 404
- Mapas citogenéticos, 406
- Mapas físicos e bancos de clones, 406
- Clonagem de genes pela posição no mapa, 408
- Caminhadas e saltos no cromossomo, 409
- Projeto genoma humano, 410
- Mapeamento do genoma humano, 410
- Sequenciamento do genoma humano, 411
- Projeto HapMap humano, 415
- Ensaios de RNA e proteína da função genômica, 416
- Sequências expressas, 416
- Microarranjos e *chips* gênicos, 417
- Proteína fluorescente verde como repórter da síntese proteica, 417
- Genômica comparativa, 421
- Bioinformática, 421
- Genomas procarióticos, 423
- Bactéria viva com genoma sintetizado quimicamente, 424
- Genomas de cloroplastos e mitocôndrias, 426
- Genomas eucarióticos, 429
- Evolução do genoma em gramíneas cereais, 430
- Evolução do genoma em mamíferos, 431

Capítulo 16

Aplicações da Genética Molecular, 439

- Uso de tecnologia do DNA recombinante para identificar genes humanos e diagnosticar doenças humanas, 440
- Doença de Huntington, 440
- Fibrose cística, 445
- Diagnóstico molecular de doenças humanas, 448
- Terapia gênica humana, 448
- Análise do perfil de DNA, 453
- Testes de paternidade, 457
- Aplicações forenses, 457
- Produção de proteínas eucarióticas em bactérias, 459

- Hormônio do crescimento humano, 459
- Proteínas com aplicação industrial, 460

Animais e vegetais transgênicos, 461

- Animais transgênicos | Microinjeção de DNA em ovos fertilizados e transfecção de células-tronco embrionárias, 461
- Vegetais transgênicos | O plasmídeo TI de *Agrobacterium tumefaciens*, 462

- Genética reversa | Análise de processos biológicos por inibição da expressão gênica, 465
- Mutações *knockout* em camundongo, 465
- Inserções de T-DNA e transpósons, 467
- Interferência por RNA, 467

Capítulo 17

Elementos Genéticos Transponíveis, 475

- Elementos transponíveis | Considerações gerais, 476
- Elementos transponíveis em bactérias, 477
- Elementos IS, 477
- Transpósons compostos, 478
- Elemento Tn3, 479
- Transpósons de "cortar e colar" em eucariotos, 480
- Elementos *Ac* e *Ds* no milho, 480
- Elementos *P* e disgenesia híbrida em *Drosophila*, 482
- Retrovírus e retrotranspósons, 485
- Retrovírus, 485
- Elementos semelhantes a retrovírus, 487
- Retrospósons, 490
- Elementos transponíveis em seres humanos, 490
- Significado genético e evolutivo dos elementos transponíveis, 492
- Transpósons como mutágenos, 492
- Transformação genética com transpósons, 493
- Transpósons e organização do genoma, 494

Capítulo 18

Regulação da Expressão Gênica em Procariotos, 500

- Expressão de genes constitutivos, induzíveis e repressíveis, 501
- Controle positivo e negativo da expressão gênica, 503
- Óperons | Unidades de expressão gênica de regulação coordenada, 505
- Óperon lactose em *E. coli* | Indução e repressão catabólica, 507
- Indução, 507
- Repressão catabólica, 512
- Interações proteína-DNA que controlam a transcrição do operon *lac*, 513

Óperon de triptofano em *E. coli* | Repressão e atenuação, 514

Repressão, 515
Atenuação, 515

Regulação da expressão gênica por controle da tradução, 520

Mecanismos reguladores pós-tradução, 521

Capítulo 19

Regulação da Expressão Gênica em Eucariotos, 526

Mecanismos de regulação da expressão gênica eucariótica | Considerações gerais, 527

Dimensões da regulação gênica eucariótica, 527

Transcrição de DNA controlada, 527

Recomposição alternativa de RNA, 528

Controle citoplasmático da estabilidade do RNA mensageiro, 528

Indução da atividade de transcrição por fatores ambientais e biológicos, 529

Temperatura | Os genes do choque térmico, 530

Moléculas sinalizadoras | Genes que respondem a hormônios, 530

Controle molecular da transcrição em eucariotos, 532

Sequências de DNA implicadas no controle da transcrição, 532

Proteínas implicadas no controle da transcrição | Fatores de transcrição, 533

Regulação pós-transcricional da expressão gênica por interferência por RNA, 536

Vias de RNAi, 536

Fontes de RNA de interferência curtos e microRNA, 536

Expressão gênica e organização da cromatina, 539

Eucromatina e heterocromatina, 539

Organização molecular do DNA transcricionalmente ativo, 540

Remodelagem da cromatina, 541

Metilação do DNA, 541

Imprinting, 542

Ativação e inativação de cromossomos inteiros, 543

Inativação de cromossomos X em mamíferos, 545

Hiperativação de cromossomos X em *Drosophila*, 546

Hipoativação de cromossomos X em *Caenorhabditis*, 546

Capítulo 20

Controle Genético do Desenvolvimento Animal, 552

Perspectiva genética sobre o desenvolvimento, 553

Atividade gênica materna no desenvolvimento, 554

Genes de efeito materno, 555

Determinação dos eixos dorsoventral e anteroposterior, 555

Atividade gênica zigótica no desenvolvimento, 558

Segmentação do corpo, 559

Formação de órgãos, 561

Especificação de tipos celulares, 563

Análise genética do desenvolvimento em vertebrados, 564

Homólogos em vertebrados de genes de invertebrados, 564

O camundongo | Mutações por inserção aleatória e mutações *knockout* gene-específicas, 565

Estudos com células-tronco de mamíferos, 566

Clonagem reprodutiva, 567

Alterações genéticas na diferenciação das células imunes de vertebrados, 568

Capítulo 21

Base Genética do Câncer, 574

Câncer | Uma doença genética, 575

As muitas formas de câncer, 575

Câncer e ciclo celular, 576

Câncer e morte celular programada, 577

Base genética do câncer, 577

Oncogenes, 578

Retrovírus indutores de tumor e oncogenes virais, 578

Homólogos celulares de oncogenes virais | Os proto-oncogenes, 579

Oncogenes celulares mutantes e câncer, 579

Rearranjos cromossômicos e câncer, 581

Genes supressores tumorais, 582

Cânceres hereditários e a hipótese de dois eventos de Knudson, 582

Papéis celulares das proteínas supressoras de tumor, 584

pRB, 584

p53, 585

pAPC, 588

pMSH2, 590

pBRCA1 e pBRCA2, 590

Vias genéticas da carcinogênese, 590

Capítulo 22

Herança de Características Complexas, 598

Características complexas, 599

Quantificação de características complexas, 599

Fatores genéticos e ambientais influenciam as características quantitativas, 599

Vários genes influenciam as características quantitativas, 599

Características de limiar, 601

Estatística em genética quantitativa, 602

Distribuições de frequência, 602

A média e a classe modal, 603

A variância e o desvio padrão, 603

Análise de características quantitativas, 604

Hipótese dos fatores múltiplos, 604

Partição da variância fenotípica, 605

Herdabilidade em sentido amplo, 605

Herdabilidade em sentido restrito, 606

Previsão de fenótipos, 607

Seleção artificial, 608

Loci de característica quantitativa, 610

Correlações entre parentes, 614

Correlação de fenótipos quantitativos entre parentes, 614

Interpretação das correlações entre parentes, 615

Genética quantitativa de características comportamentais humanas, 616

Inteligência, 617

Personalidade, 617

Capítulo 23

Genética de Populações, 622

Teoria das frequências alélicas, 623

Estimativa das frequências alélicas, 623

Relação entre frequências genotípicas e frequências alélicas | O princípio de Hardy-Weinberg, 623

Aplicações do princípio de Hardy-Weinberg, 624

Exceções ao princípio de Hardy-Weinberg, 625

Uso das frequências alélicas no aconselhamento genético, 627

Seleção natural, 628

Conceito de aptidão, 628

Seleção natural no nível do gene, 628

Deriva genética aleatória, 631

Alterações aleatórias das frequências alélicas, 631

Efeitos do tamanho da população, 632

Populações em equilíbrio genético, 633

Seleção balanceadora, 633

Balanço mutação-seleção, 635

Balanço mutação-deriva, 635

Capítulo 24

Genética Evolutiva, 642

Surgimento da teoria evolutiva, 643

Teoria da evolução de Darwin, 643

Genética evolutiva, 644

Variação genética em populações naturais, 644

Variação dos fenótipos, 644

Variação da estrutura cromossômica, 646

Variação da estrutura das proteínas, 646

Variação das sequências nucleotídicas, 646

Evolução molecular, 648

Moléculas como “documentos da história evolutiva”, 648

Filogenias moleculares, 649

Taxas de evolução molecular, 650

Relógio molecular, 652

Variação na evolução das sequências de proteínas, 653

Variação na evolução das sequências de DNA, 653

Teoria neutra da evolução molecular, 654

Evolução molecular e evolução fenotípica, 655

Especiação, 657

O que é uma espécie?, 657

Modos de especiação, 658

Evolução humana, 660

Seres humanos e grandes primatas, 660

Evolução humana no registro fóssil, 661

Variação da sequência de DNA e origens humanas, 661

Apêndice A | Regras da Probabilidade, 669

Apêndice B | Probabilidades Binomiais, 670

Apêndice C | Hibridização *in Situ*, 671

Apêndice D | Evidências de um RNA Mensageiro Instável, 673

Apêndice E | Taxas Evolutivas, 675

Respostas dos Problemas de Número Ímpar, 677

Glossário, 703

Créditos das Fotos, 728

Créditos das Ilustrações, 730

Índice Alfabético, 731

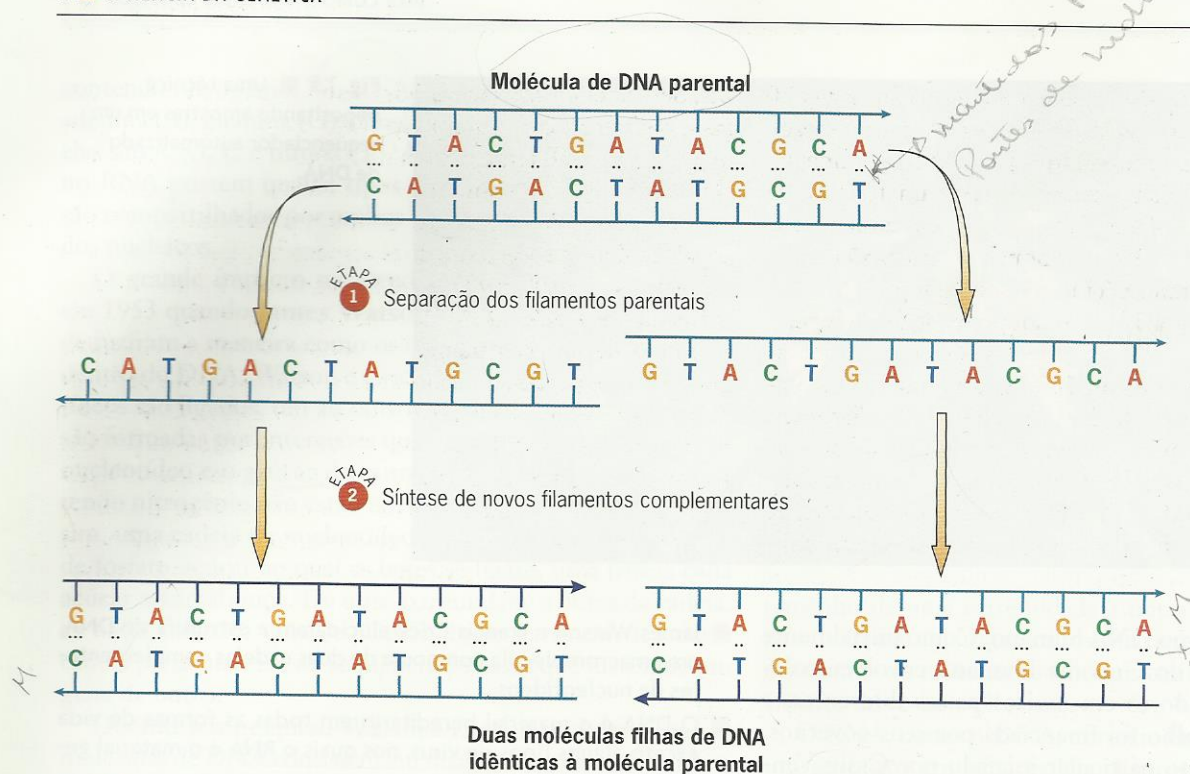


Fig. 1.6 ■ Replicação do DNA. Separação de filamentos (em cima) e a síntese de filamentos novos complementares (embaixo) ocorrem de modo coordenado e processual durante replicação do DNA. Aqui eles estão separados para destacar estes dois aspectos do processo geral.

O processo de replicação do DNA é baseado na natureza complementar dos filamentos que constituem as moléculas bifilamentares de DNA (Fig. 1.6). Esses filamentos são mantidos juntos por pontes de hidrogênio entre pares de bases específicos – A pareada com T e G pareada com C. Quando tais ligações são quebradas, os filamentos separados podem servir como moldes para a síntese de novos filamentos complementares. Os novos filamentos são montados pela incorporação gradativa de nucleotídeos opostos a nucleotídeos nos filamentos moldes. Essa incorporação está de acordo com as regras de pareamento de bases. Assim, a sequência de nucleotídeos em um filamento que está sendo sintetizado é ditada pela sequência de nucleotídeos do filamento molde. No final do processo de replicação, cada filamento molde é pareado com um novo filamento complementar sintetizado. Assim, são criados dois DNA bifilamentares idênticos a partir da dupla hélice original.

O processo de replicação do DNA não ocorre espontaneamente. Como a maioria dos processos bioquímicos, ela é catalisada por enzimas. Exploraremos os detalhes da replicação do DNA, incluindo os papéis desempenhados por diferentes enzimas, no Cap. 10.

Expressão Gênica: Uso da Informação Genética

As moléculas de DNA contêm informações para dirigir as atividades das células e guiar o desenvolvimento, o funcionamento e o comportamento dos organismos que são

feitos destas células. Esta informação é codificada em seqüências de nucleotídeos dentro das moléculas de DNA do genoma. Entre os organismos celulares, o menor genoma conhecido é o de *Mycoplasma genitalium*: 580.070 pares de nucleotídeos. Em contraste, o genoma humano consiste em 3,2 bilhões de pares de nucleotídeos. Nestes e em todos os outros genomas, a informação contida dentro do DNA é organizada nas unidades que chamamos de genes. Um *M. genitalium* possui 482 genes, enquanto um espermatozóide humano possui de 20.000 a 25.000. Cada gene é um trecho de pares de nucleotídeos ao longo de uma molécula de DNA. Uma determinada molécula de DNA pode conter milhares de genes diferentes. Em uma célula de *M. genitalium*, todos os genes estão situados em uma molécula de DNA – o único cromossomo deste organismo. Em um espermatozóide humano, os genes estão situados em 23 moléculas diferentes de DNA correspondentes aos 23 cromossomos da célula. A maioria do DNA em *M. genitalium* compreende genes, enquanto a maioria do DNA humano não, isto é, a maioria do DNA humano não é constituída de genes. Vamos investigar a composição gênica e não-gênica dos genomas em muitas partes deste livro, especialmente no Cap. 16.

Como a informação dentro de genes individuais é organizada e expressa? Esta pergunta é central na genética, e dedicaremos nossa atenção a ela nos Caps. 11, 12 e 14. Aqui, é suficiente dizer que a maioria dos genes contém as instruções para a síntese de proteínas. Cada proteína consiste em um ou

mais polipeptídeos, que são cadeias de aminoácidos. Os 20 tipos diferentes de aminoácidos que ocorrem naturalmente podem ser combinados de mil modos para formar polipeptídeos. Cada polipeptídeo tem uma sequência característica de aminoácidos. Alguns polipeptídeos são curtos, com apenas alguns aminoácidos de tamanho, enquanto outros são enormes, com milhares de aminoácidos de tamanho.

A sequência de aminoácidos em um polipeptídeo é especificada por uma sequência de unidades codificantes elementares dentro de um gene. Essas unidades codificantes elementares, chamadas **códons**, são trincas de nucleotídeos adjacentes. Um gene típico pode conter centenas ou mesmo milhares de códons. Cada códon especifica a incorporação de um aminoácido em um polipeptídeo. Assim, a informação codificada dentro de um gene é usada para dirigir a síntese de um polipeptídeo, que em geral é chamado de produto do gene. Às ve-

zes, dependendo de como a informação genética é usada, um gene pode codificar vários polipeptídeos; entretanto, esses polipeptídeos geralmente são todos correlatos, compartilhando algumas sequências comuns de aminoácidos.

A expressão da informação genética para formar um polipeptídeo é um processo de duas etapas (Fig. 1.7). Primeira, a informação contida no DNA de um gene é copiada em uma molécula de RNA. O RNA é montado de modo gradativo ao longo de um dos filamentos da dupla hélice do DNA. Durante este processo de montagem, A no RNA faz par com T no DNA, G no RNA faz par com C no DNA, C no RNA faz par com G no DNA e U no RNA faz par com A no DNA. Assim, a sequência de nucleotídeos do RNA é determinada pela sequência de nucleotídeos de um filamento de DNA no gene. O processo que produz esta molécula de RNA é chamado **transcrição**, e o próprio RNA é chamado **trans-**

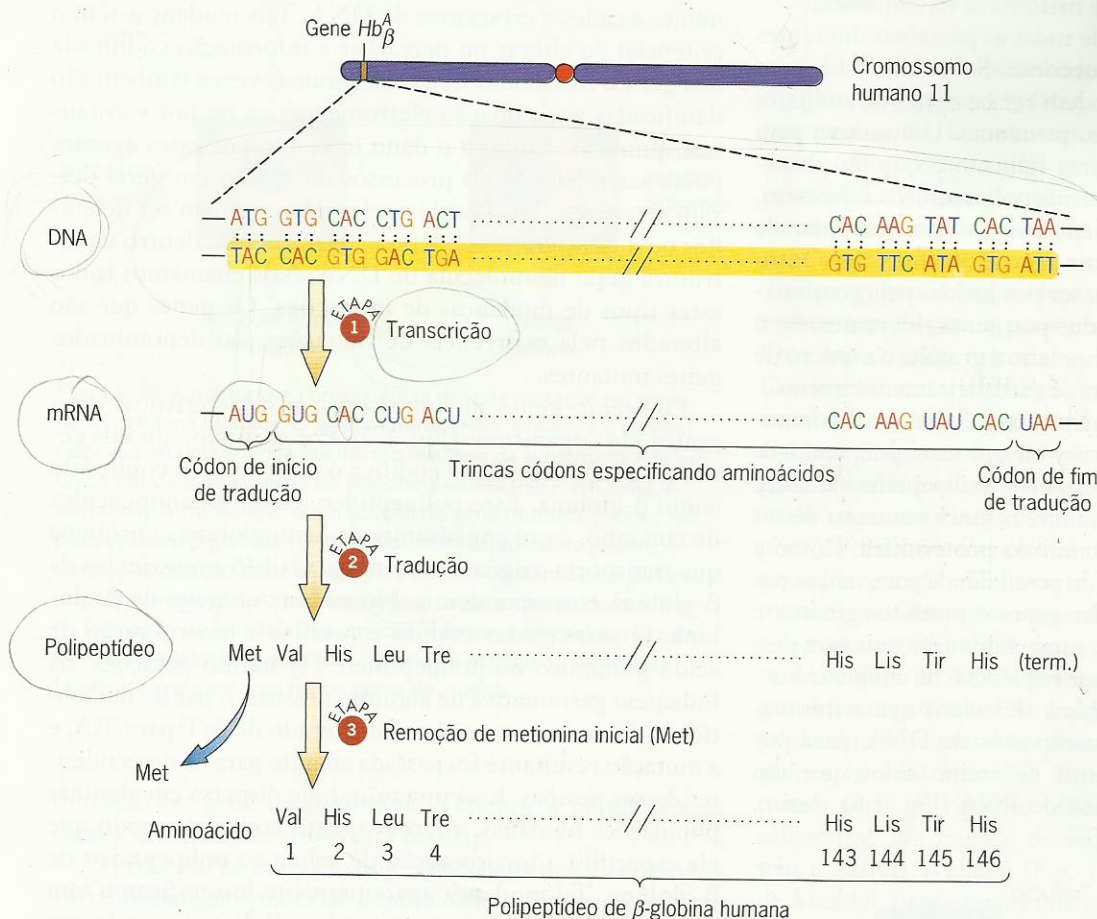


Fig. 1.7 ■ Transcrição e tradução, as etapas chaves na expressão da informação genética, são ilustradas usando-se a síntese de β -globina humana, um dos polipeptídeos constituintes da hemoglobina. Durante a transcrição (etapa 1), os dois filamentos do DNA estão localmente deslicoidizados, e um dos filamentos, o inferior mostrado aqui em destaque, serve como molde para a síntese de um filamento complementar de RNA. Após sofrer modificações, o mRNA (RNA mensageiro) resultante é usado como molde para a síntese do polipeptídeo de β -globina pelo processo de tradução (etapa 2). Durante a tradução, cada trinca códon no mRNA especifica a incorporação de um aminoácido na cadeia polipeptídica. A tradução é iniciada por um códon de começo, que especifica a incorporação do aminoácido metionina (met), e é terminada por um códon de fim, que não especifica a incorporação de nenhum aminoácido. Após a tradução estar completa, a metionina inicial é removida (etapa 3) para produzir o polipeptídeo β -globina final.

critico. O RNA transcrito acaba por se separar de seu DNA molde e, em alguns organismos, é alterado pela adição, pela deleção ou pela modificação de nucleotídeos. A molécula terminada, chamada **RNA mensageiro** ou simplesmente **mRNA**, contém toda a informação necessária para a síntese de um polipeptídeo.

O segundo estágio na expressão da informação de um gene é denominado **tradução**. Neste estágio, o mRNA de um gene atua como um molde para a síntese de um polipeptídeo. Cada um dos códons do gene, agora presentes na sequência do mRNA, especifica a incorporação de um determinado aminoácido na cadeia polipeptídica. Um aminoácido é adicionado de cada vez. Assim, o polipeptídeo é produzido gradativamente lendo os códons de uma ponta a outra do mRNA. Quando o polipeptídeo está pronto, ele se dissocia do mRNA e desempenha seu papel na célula. Alguns polipeptídeos são alterados pela remoção do primeiro aminoácido, que geralmente é metionina, na sequência.

Nós chamamos a coleção de todas as proteínas diferentes em um organismo de seu **proteoma**. Seres humanos, com de 20.000 a 25.000 genes, podem ter centenas de milhares de proteínas diferentes em seu proteoma. Um motivo para o grande tamanho do proteoma humano é que um determinado gene pode codificar vários polipeptídeos diferentes, mas correlatos, e esses polipeptídeos podem combinar de modos complexos para produzir proteínas diferentes. Outro motivo é que proteínas podem ser produzidas pela combinação de polipeptídeos codificados por genes diferentes. Se o número de genes no genoma humano é grande, o número de proteínas no proteoma humano é verdadeiramente imenso.

O estudo de todas as proteínas nas células – sua composição, as sequências de aminoácidos em seus polipeptídeos constituintes, as interações entre esses polipeptídeos e entre proteínas diferentes e, logicamente, o funcionamento dessas moléculas complexas – é denominado **proteômica**. Como a genômica, a proteômica tem sido possibilitada por avanços nas tecnologias usadas para estudar genes e produtos gênicos e pelo desenvolvimento de programas computacionais para pesquisar bancos de dados e analisar sequências de aminoácidos.

Por todas estas considerações, fica claro que a informação flui dos genes, que são compostos de DNA, para polipeptídeos, que são compostos de aminoácidos, por um intermediário, que é composto de RNA (Fig. 1.8). Assim,

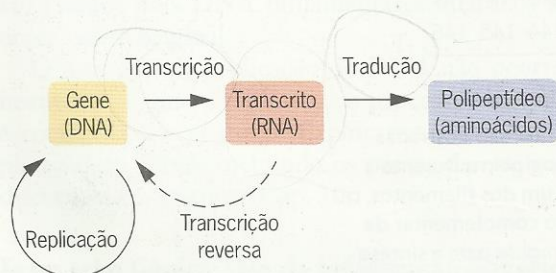


Fig. 1.8 ■ O dogma central da biologia molecular mostrando como a informação genética é propagada (por replicação do DNA) e expressa (por transcrição e tradução). Na transcrição reversa, o RNA é usado como um molde para a síntese de DNA.

em sentido amplo, o fluxo de informações é DNA→RNA→polipeptídeo, uma progressão geralmente chamada **dogma central da biologia molecular**. Em vários capítulos, veremos circunstâncias nas quais a primeira parte desta progressão é revertida, isto é, o RNA é usado como um molde para a síntese de DNA. Este processo é denominado **transcrição reversa**, desempenha um papel importante nas atividades de alguns tipos de vírus, incluindo o vírus que causa a síndrome da imunodeficiência adquirida, ou AIDS. Também afeta profundamente o conteúdo e a estrutura dos genomas de muitos organismos, incluindo o genoma humano.

Mutação: Mudança na Informação Genética

A replicação do DNA é um processo extremamente preciso, mas não é perfeito. Com uma frequência baixa porém mensurável, nucleotídeos são incorporados incorretamente a cadeias crescentes de DNA. Tais mudanças têm o potencial de alterar ou perturbar a informação codificada nos genes. Moléculas de DNA algumas vezes também são danificadas por radiação eletromagnética ou por substâncias químicas. Embora o dano induzido por esses agentes possa ser reparado, os processos de reparo em geral deixam cicatrizes. Trechos de nucleotídeos podem ser deletados ou duplicados ou podem ser rearranjados dentro da estrutura geral da molécula de DNA. Nós chamamos todos estes tipos de mudanças de **mutações**. Os genes que são alterados pela ocorrência de mutações são denominados genes mutantes.

Em geral, genes mutantes causam características diferentes nos organismos (Fig. 1.9). Por exemplo, um dos genes do genoma humano codifica o polipeptídeo conhecido como β -globina. Este polipeptídeo, com 146 aminoácidos de tamanho, é um constituinte da hemoglobina, a proteína que transporta oxigênio no sangue. Os 146 aminoácidos da β -globina correspondem a 146 códons no gene de β -globina. O sexto destes códons especifica a incorporação de ácido glutâmico no polipeptídeo. Há muitas gerações, na linhagem germinativa de algumas pessoas, o par de nucleotídeos do meio neste códon foi alterado de A:T para T:A, e a mutação resultante foi passada adiante para os descendentes dessas pessoas. Essa mutação, hoje dispersa em algumas populações humanas, alterou o sexto códon de modo que ele especifica a incorporação de valina ao polipeptídeo de β -globina. Tal mudança aparentemente insignificante tem um efeito deletério na estrutura das células que produzem e estocam hemoglobina, as hemácias. Pessoas que levam apenas a versão mutante do gene de β -globina têm hemácias falcêmicas, enquanto pessoas que possuem apenas a versão não mutante deste gene têm hemácias em forma de disco. As células falcêmicas não transportam oxigênio eficientemente pelo corpo. Consequentemente, pessoas com hemácias falciformes desenvolvem uma doença grave, tão grave que de fato podem morrer dela. Essa doença, chamada anemia falciforme, é portanto relacionada com uma

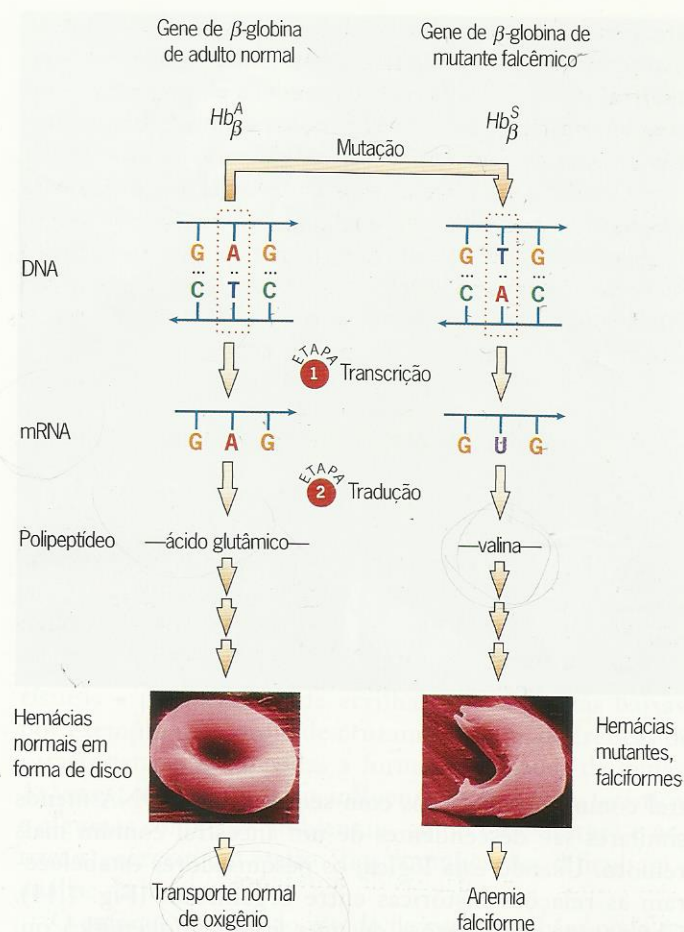


Fig. 1.9 ■ A natureza e consequ ncia de uma muta  o no gene para β -globina humana. O gene mutante (Hb^S ; em cima,   direita) respons vel pela anemia falciforme resultou da substitui  o de um  nico par de base no gene de β -globina adulta (Hb^A ; em cima,   esquerda). A transcri  o e a tradu  o do gene mutante produz um polipept deo β -globina contendo o amino cido valina (centro,   direita) na posi  o onde a β -globina normal cont m  cido glut mico (centro,   esquerda). A mudan a de um  nico amino cido resulta na forma  o de hem cias falc micas (inferior,   direita) em vez de c lulas normalmente disc ides (inferior,   esquerda). As c lulas falc micas causam uma forma grave de anemia.

muta  o no gene de β -globina. Veremos a natureza e as causas de muta  es como esta no Cap. 13.

O processo de muta  o tem outro aspecto – ele introduz variabilidade no material gen tico dos organismos. Com o tempo, os genes mutantes criados por muta  o podem espalhar-se em uma popula  o, como o gene mutante de β -globina se espalhou em algumas popula  es humanas. Tais mudan as na composi  o gen tica de uma popula  o constituem a base da evolu  o biol gica.

PONTOS IMPORTANTES

- Quando o DNA se replica, cada filamento da dupla h lice serve como um molde para a s ntese de um filamento complementar.

- Quando a informa  o gen tica   expressa, um filamento do DNA d plex de um gene   usado como molde para a s ntese de um filamento complementar de RNA.
- Para a maioria dos genes, a s ntese de RNA (transcri  o) gera uma mol cula (o RNA transcrito) que se torna um RNA mensageiro (mRNA).
- A informa  o codificada em um mRNA   traduzida em uma seq  ncia de amino cidos em um polipept deo.
- Muta  es podem alterar a seq  ncia de DNA de um gene.
- A variabilidade gen tica criada por muta  o   a base da evolu  o biol gica.

GEN TICA E EVOLU  O

A gen tica tem muito a contribuir para o estudo cient fico da evolu  o.

O registro da hist ria humana evolutiva perfez alguns milhares de anos. Artefatos, como pinturas em cavernas, pontas de setas, ferramentas de pedra e ossos, e outros materiais, d o-nos uma id ia mais profunda da vida humana no passado. Em algum ponto, entretanto, at  mesmo a evid ncia arqueol gica da atividade humana desaparece, e ficamos com a pergunta “Quem somos n s?” e “De onde viemos?” Tais perguntas sobre a identidade e a origem dos seres humanos nos for am a pensar sobre nossa heran a gen tica. Que aspectos do material gen tico nos tornam humanos? Como esse material chegou at  n s vindo de nossos ancestrais? Ele mudou com o tempo? Nosso material gen tico   relacionado como o de outras esp cies? Nossa hist ria   parte de uma hist ria maior da vida?

As respostas a essas perguntas devem levar em conta a replica  o mais ou menos fiel do DNA e sua transmiss o de gera  o a gera  o. Entretanto, ocorrem muta  es, e os genes mutantes resultantes s o copiados e transmitidos para gera  es subsequentes. Gradualmente, ao logo do tempo, muta  es se acumulam no DNA, e vemos seus efeitos como diferen as entre organismos. As linhagens mendelianas de ervilhas possuíam genes mutantes diferentes, e tamb m as pessoas de grupos  tnicos diferentes. Em quase todas as esp cies, pelo menos parte da varia  o observ vel tem uma base gen tica subjacente. Na metade do s culo dezenove, **Charles Darwin** e **Alfred Wallace** (Fig. 1.10), ambos contempor neos de Mendel, propuseram que essa varia  o possibilita que as esp cies mudem – isto  , evoluam – com o tempo.

As id ias de Darwin e Wallace revolucionaram o pensamento cient fico. Elas introduziram uma perspectiva hist rica na biologia e deram credibilidade ao conceito de que todos os seres vivos s o relacionados por descend ncia de um ancestral comum. Entretanto, quando tais id ias foram propostas, o trabalho de Mendel sobre a hereditariedade ainda estava em progresso, e a ci ncia da gen tica n o tinha nascido. A pesquisa na evolu  o biol gica foi estimulada quando as descobertas de Mendel vieram   luz

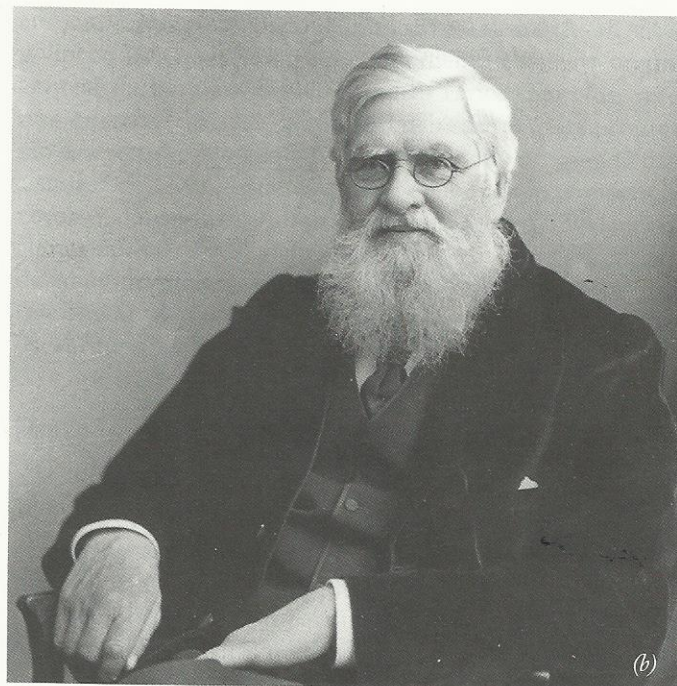
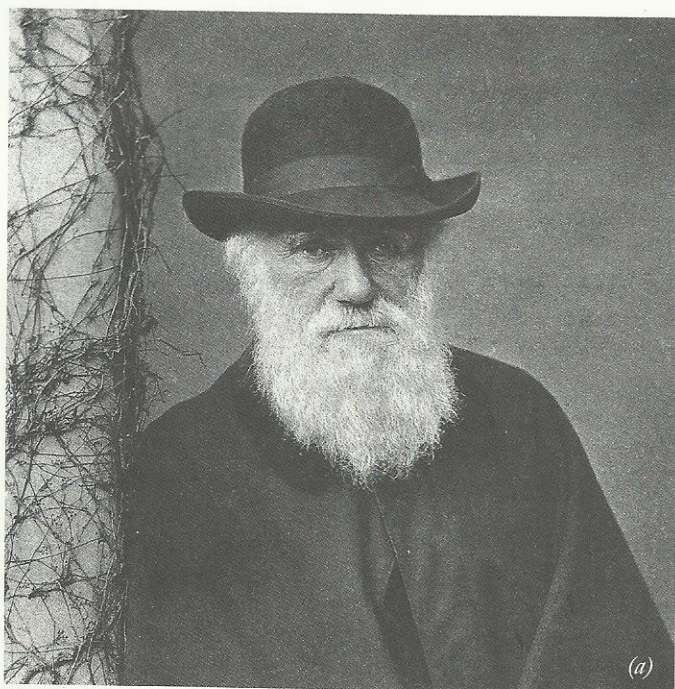


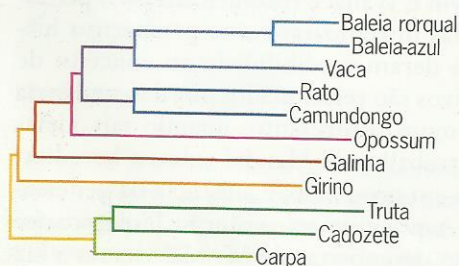
Fig. 1.10 ■ (a) Charles Darwin. (b) Alfred Wallace.

no começo do século vinte e tiveram novo avanço quando surgiram as técnicas de sequenciamento do DNA, no final do século. Com o sequenciamento do DNA, podemos ver similaridades e diferenças no material genético de vários organismos. Na suposição de que as sequências de nucleotídeos no DNA são o resultado de processos históricos, é possível interpretar tais similaridades e diferenças em função do tempo. Organismos com sequências de DNA muito similares são descendentes de um recente ance-

tral comum; organismos com sequências de DNA menos similares são descendentes de um ancestral comum mais remoto. Usando esta lógica, os pesquisadores estabeleceram as relações históricas entre organismos (Fig. 1.11). Chamamos tais relações de uma árvore filogenética ou, mais simplesmente, uma **filogenia**, das palavras gregas que significam “a origem das tribos”.

Hoje em dia, a construção de árvores filogenéticas é parte importante do estudo da evolução. Biólogos usam os dados da sequência do DNA dos projetos do genoma e outras pesquisas, tais como o programa “Tree of Life” do United States National Science Foundation, em combinação com dados anatômicos coletados de organismos vivos e fósseis para distinguir as relações evolutivas entre espécies. Exploraremos a base genética da evolução nos Caps. 26 e 27.

Fig. 1.11 ■ Árvore filogenética mostrando as relações evolutivas entre 11 vertebrados diferentes. Esta árvore foi construída comparando-se as sequências do gene para citocromo b, que é levado pelo DNA encontrado nas mitocôndrias desses animais. Os 11 animais diferentes foram posicionados na árvore de acordo com a similaridade de suas sequências do gene de citocromo b. Esta árvore é consistente com outras informações (ex., dados obtidos pelo estudo de fósseis), exceto no caso das posições de três espécies de peixes. O cadozete é de fato mais proximamente relacionado à carpa do que à truta. Essa discrepância indica a necessidade de interpretar os resultados das comparações de sequências de DNA com cuidado.



PONTOS IMPORTANTES

- A evolução depende da ocorrência, da transmissão e da dispersão de genes mutantes em grupos de organismos.
- A sequência de dados do DNA fornece um modo para estudar o processo histórico da evolução.

NÍVEIS DE ANÁLISE GENÉTICA

Os geneticistas abordam sua ciência por diferentes pontos de vista – de um gene, uma molécula de DNA ou uma população de organismos.

A análise genética é praticada em níveis diferentes. O tipo mais antigo de análise genética segue os passos de

Mendel enfocando como as características são herdadas quando linhagens diferentes de organismos são hibridizadas. Um outro tipo de análise genética segue os passos de Watson e Crick e as pessoas que trabalharam em vários projetos de genoma enfocando a constituição molecular do material genético. Outro tipo ainda de análise genética imita Darwin e Wallace enfocando populações inteiras de organismos. Todos estes níveis de análise genética são rotineiramente usados em pesquisa hoje em dia. Embora os encontremos em locais diferentes deste livro, daremos uma breve descrição deles aqui.

Genética Clássica

O período anterior à descoberta da estrutura do DNA em geral é chamado de era da **genética clássica**. Durante essa época, geneticistas abordavam sua ciência analisando os resultados de cruzamentos entre linhagens diferentes de organismos, como Mendel havia feito em seu trabalho com ervilhas. Neste tipo de análise, os genes são identificados estudando-se a herança de diferenças de características – plantas altas de ervilhas *versus* plantas baixas, por exemplo – na prole de cruzamentos. As diferenças de características são devidas a formas alternadas de genes. Às vezes, mais de um gene influenciam uma característica, e algumas condições ambientais, como temperatura e nutrição, exercem um efeito. Tais complicações dificultam a análise da herança.

O enfoque clássico do estudo de genes também pode ser coordenado com estudos da estrutura e do comportamento dos cromossomos, que são as entidades celulares que contêm os genes. Analisando os padrões de herança, os geneticistas podem localizar genes em cromossomos específicos. Uma análise mais detalhada permite-lhes localizar os genes em posições específicas dentro dos cromossomos, uma prática chamada mapeamento cromossômico. Como esses estudos enfatizam a transmissão de genes e cromossomos de uma geração para a seguinte, em geral eles são chamados exercícios na **genética de transmissão**. Entretanto, a genética clássica não é limitada à análise de transmissão de genes e cromossomos. Ela também estuda a natureza do material genético – como ele controla as características e como ele sofre mutação. Apresentamos as características da genética clássica nos Caps. 3 a 8.

Genética Molecular

Com a descoberta da estrutura do DNA, a genética entrou em uma nova fase. A replicação, a expressão e a mutação dos genes hoje podem ser estudadas no nível molecular. Este enfoque à análise genética passou para um novo nível quando se tornou possível seqüenciar facilmente moléculas de DNA. A análise genética molecular é baseada no estudo de seqüências de DNA. O conhecimento de uma seqüência de DNA e as comparações com outras seqüências de DNA permitem que um geneticista

defina um gene quimicamente. Os componentes internos – seqüências codificantes, seqüências reguladoras e seqüências não codificantes – podem ser identificados, e a natureza do polipeptídeo codificado pelo gene pode ser prevista.

Mas o enfoque molecular à análise genética é muito mais que o estudo de seqüências de DNA. Os geneticistas aprenderam a cortar moléculas de DNA em pontos específicos. Genes inteiros, ou trechos de genes, podem ser removidos de uma molécula de DNA e inseridos em outra molécula de DNA. Essas moléculas de DNA “recombinantes” podem ser replicadas em bactérias ou mesmo em tubos de ensaio que receberam enzimas apropriadas. Miligramas de um determinado gene podem ser geradas no laboratório em uma tarde. Em resumo, os geneticistas aprenderam a manipular genes mais ou menos à vontade. Essa manipulação permitiu que pesquisadores estudem fenômenos genéticos em grande detalhe. Eles aprenderam até a transferir genes de um organismo para outro. Apresentamos exemplos de análise de genética molecular em muitos dos capítulos deste livro.

Genética de Populações

A genética também pode ser estudada em populações inteiras de organismos. Os indivíduos dentro de uma população podem levar alelos diferentes de um gene; talvez eles levem alelos diferentes de muitos genes. Essas diferenças tornam os indivíduos distintos do ponto de vista genético, possivelmente mesmo únicos. Em outras palavras, os membros de uma população variam em sua constituição genética. Os geneticistas procuram documentar essa variabilidade e compreender seu significado. Seu enfoque mais básico é determinar as freqüências de alelos específicos em uma população e então avaliar se tais freqüências mudam com o tempo. Caso sim, a população está evoluindo. A avaliação da variabilidade genética em uma população é portanto uma base para o estudo da evolução biológica. Também é útil no esforço para compreender a herança de características complexas, tais como tamanho do corpo ou suscetibilidade a doença. Em geral, as características complexas são de considerável interesse porque têm um significado do ponto de vista agrícola ou médico. Nós discutiremos a análise genética em nível populacional nos Caps. 25, 26 e 27.

PONTOS IMPORTANTES

- Na análise genética clássica, os genes são estudados seguindo a herança de características em cruzamentos entre linhagens diferentes de um organismo.
- Na análise genética molecular, os genes são estudados mediante isolamento, seqüenciamento e manipulação do DNA e examinando-se os produtos da expressão gênica.
- Na análise genética de populações, os genes são estudados avaliando-se a variabilidade entre indivíduos em um grupo de organismos.

A GENÉTICA NO MUNDO: APLICAÇÕES DA GENÉTICA A ESFORÇOS HUMANOS

A genética é relevante em muitos aspectos fora do laboratório de pesquisa.

A genética começou em um monastério europeu; hoje, é um empreendimento mundial. O significado e o escopo internacional da genética são evidentes nas publicações científicas de hoje, que destacam o trabalho de geneticistas de muitos países diferentes. Eles também são evidentes na miríade de modos pelos quais a genética é aplicada à agricultura, à medicina e a muitos outros interesses humanos em todo o mundo. Consideraremos algumas destas aplicações nos Caps. 15, 16, 17, 24 e 25. Alguns dos destaques são introduzidos adiante.

Genética na Agricultura

Na época em que surgiram as primeiras civilizações, os seres humanos já tinham aprendido a cultivar plantas e criar rebanhos. Eles também tinham aprendido a melhorar seus plantios e criações de animais por cruzamento seletivo. Esta aplicação pré-mendeliana dos princípios genéticos teve poderosos efeitos. Por milhares de gerações, espécies de plantas e animais domesticados tornaram-se muito diferentes de seus ancestrais selvagens. Por exemplo, o gado mudou de aspecto e comportamento (Fig. 1.12), e o milho, que é descendente de uma gramínea chamada teosinto (Fig. 1.13), mudou tanto que não pode mais desenvolver-se sem o cultivo humano.

Os programas de cruzamentos seletivos, hoje baseados na teoria genética, continuam a ter papéis importantes na agricultura. Variedades com alta produtividade de trigo, milho, arroz e muitas outras plantas foram desenvolvidas por criadores para alimentar a população humana crescente. Técnicas de reprodução seletiva também foram aplicadas a animais tais como gado de corte e leiteiro, suínos e ovelhas e a plantas hortícolas tais como árvores, gramas e flores de jardim.

Começando nos anos 1980, os enfoques clássicos de cultivo e criação foram suplementados, e em alguns casos suplantados, por enfoques de genética molecular. Foram construídos mapas genéticos detalhados de cromossomos de várias espécies para destacar genes de significado na agricultura. Localizando genes para características tais como pro-

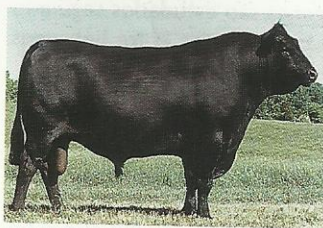


Fig. 1.13 ■ Espigas de milho (direita) e seu ancestral, teosinto (esquerda).

dução de grãos ou resistência a doenças, os criadores agora podem criar esquemas para incorporar determinados alelos a variedades na agricultura. Tais projetos de mapeamento foram feitos intensamente e para uma espécie, o arroz, culminaram no seqüenciamento completo do genoma. Outros cultivos e projetos de seqüenciamento genômico ainda estão em andamento. Todos os tipos de genes potencialmente úteis estão sendo identificados e estudados nesses projetos.

Criadores de plantas e animais também estão empregando as técnicas de genética molecular para introduzir genes de outras espécies no cultivo de plantas e criação de animais. Este processo de mudança da constituição genética de um organismo foi inicialmente desenvolvido usando-se espécies de teste tais como as moscas-das-frutas. Hoje, é amplamente usado para aumentar o material genético de muitos tipos de criaturas. Plantas e animais que foram alterados pela introdução de genes exógenos são chamados **GMO, organismos geneticamente modificados**. O milho BT é um exemplo. Muitas variedades de milho que crescem hoje nos USA levam um gene da bactéria *Bacillus thuringiensis*. Este gene codifica uma proteína que é tóxica para muitos insetos. As linhagens de milho que levam o gene para toxina BT são resistentes a ataques pela broca do milho europeu, um inseto que causou

Fig. 1.12 ■ Raças de gado de corte.



Angus



Beefmaster



Simmental



Charolais

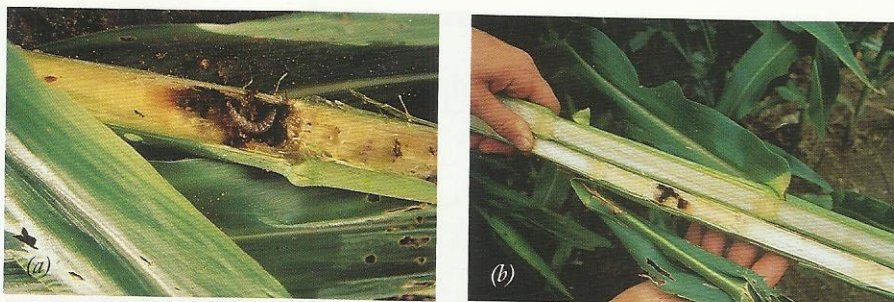


Fig. 1.14 ■ (a) Broca do milho europeu comendo o pedicelo de uma planta de milho. (b) Comparação lado a lado de pedicelos de milho de plantas que são resistentes (em cima) e suscetíveis (em baixo) à broca do milho. A planta resistente expressa um gene para uma proteína inseticida derivada de *Bacillus thuringiensis*.

um enorme dano no passado (Fig. 1.14). Assim, as plantas de milho BT produzem seu próprio inseticida.

O desenvolvimento e o uso de GMO levantou uma controvérsia mundial. Por exemplo, países africanos e europeus têm relutado em cultivar milho BT ou comprar milho BT cultivado nos EUA. Sua relutância é devida a vários fatores, incluindo o conflito de interesses de pequenos fazendeiros e grandes corporações de agricultura, bem como preocupações quanto à segurança de consumir alimentos geneticamente modificados. Há também uma preocupação de que o milho BT possa matar espécies inofensivas de insetos tais como borboletas e abelhas. Avanços na genética molecular têm fornecido as ferramentas e os materiais para mudar profundamente a agricultura. Hoje em dia, os que elaboram as políticas estão lidando com as implicações dessas novas tecnologias.

Genética na Medicina

A genética clássica forneceu aos médicos uma longa lista de doenças que são causadas por genes mutantes. O estudo dessas doenças começou logo após a redescoberta do trabalho de Mendel. Em 1909, Sir Archibald Garrod, um médico e bioquímico inglês, publicou um livro chamado *Erros Inatos do Metabolismo*. Nesse livro, Garrod documentou a maneira como anomalias metabólicas podem estar associadas a alelos mutantes. Sua pesquisa foi fundamental, e nas décadas seguintes foram identificados e catalogados grandes números de distúrbios herdados. A partir desse trabalho, os médicos aprenderam a diagnosticar doenças genéticas, a segui-las nas famílias e a prever as chances de que determinadas pessoas possam herdá-las. Hoje em dia muitos hospitais e clínicas têm profissionais conhecidos como **consultores genéticos** que são treinados para orientar pessoas quanto aos riscos de herdar ou transmitir doenças genéticas. Discutiremos alguns aspectos da consulta genética no Cap. 3.

Doenças genéticas como as que Garrod estudou são bastante raras na maioria das populações. Por exemplo, entre neonatos, a incidência de fenilcetonúria, um distúrbio do metabolismo de aminoácidos, é de apenas um em 10.000. Entretanto, genes mutantes também contribuem para doenças humanas mais prevalentes – doenças cardíacas e câncer, por exemplo. No Cap. 25, exploraremos modos de avaliar riscos genéticos para características complexas tais como suscetibilidade a doença cardíaca e no Cap. 24 investigaremos a base genética do câncer.

Os avanços em genética molecular estão fornecendo novos modos de detectar genes mutantes em pessoas. Testes diagnósticos baseados em análise do DNA hoje estão prontamente disponíveis. Por exemplo, um laboratório hospitalar pode testar uma amostra de sangue ou um *swab* de bochecha quanto à presença de um alelo mutante do gene *BRCA1*, que predispõe fortemente seus possuidores a desenvolver câncer de mama. Se uma mulher tiver o alelo mutante, ela pode ser alertada a fazer mastectomia para evitar que ocorra câncer de mama. Ela também pode ser alertada a não ter filhos, pois cada filho teria um risco de 50% de herdar o alelo mutante *BRCA1*. A aplicação destas novas tecnologias de genética molecular levanta, portanto, questões difíceis para as pessoas envolvidas.

A genética molecular também está fornecendo novos meios de tratar doenças. Por décadas, diabéticos receberam insulina obtida de animais, geralmente porcos. Hoje, a insulina humana perfeita é produzida em bactérias que possuem o gene da insulina humana. Tanques destas células produzem o polipeptídeo da insulina em escala industrial (Fig. 1.15). O hormônio de



Fig. 1.15 ■ Foto de um fermentador cultivando bactérias em larga escala.



UM MARCO NA GENÉTICA: Φ X174, o Primeiro Genoma Sequenciado do DNA

Às vezes, a vida é simples. Um vírus é uma forma de vida que consiste em um pouco mais que um ácido nucleico, seja RNA ou DNA, e alguns tipos de proteínas. O ácido nucleico, que é encapsulado por um subgrupo de proteínas, é o genoma do vírus. Ele contém genes que são necessários para formação e propagação do vírus. A maioria dos genomas virais são pequenos, e o do vírus conhecido como Φ X174 é um dos menores.

O Φ X174 infecta *Escherichia coli*, uma bactéria que vive naturalmente no intestino humano. O genoma do Φ X174 consiste em DNA, mas, ao contrário do DNA de *E. coli* e seu hospedeiro humano, o Φ X174 é unifilamentar. Esta molécula unifilamentar pode agir como um molde para a síntese de um filamento complementar. Ela o faz após entrar em uma *E. coli*, e também o faz *in vitro*, isto é, em tubo de ensaio, desde que um pesquisador forneça cada um dos quatro nucleotídeos, uma enzima que sintetiza DNA e um pequeno filamento de DNA que seja complementar a um segmento do DNA do Φ X174. Esse curto filamento de DNA serve como *primer* para iniciar o processo de síntese de DNA no molde de Φ X174.

Nos anos 1970, químicos e bioquímicos tentaram desenvolver meios para sequenciar genomas virais. Foram empregadas muitas estratégias diferentes, e o genoma do Φ X174 tornou-se um caso de teste. Por vários anos, Frederick Sanger (Fig. 1) e seus colegas do MRC Laboratory of Molecular Biology na Cambridge University no Reino Unido trabalharam no problema de sequenciamento do genoma do Φ X174. Seu enfoque foi sintetizar DNA *in vitro* usando DNA viral como molde e então analisar os produtos quanto à posição de nucleotídeos individuais. Foi um trabalho árduo. Em fevereiro de 1977, eles publicaram uma sequência quase completa de 5.375 nucleotídeos com base em seus resultados com uma técnica chamada “método do mais e menos”, pois envolvia adicionar ou subtrair nucleotídeos individuais de uma mistura de reação *in vitro*. Em uma reflexão sobre seu trabalho, Sanger e seus colaboradores escreveram:

Como em outros métodos de sequenciamento de ácidos nucleicos, a técnica de mais e menos usada por si só não pode ser vista como um sistema totalmente confiável, e erros ocasionais podem ocorrer. Tais erros e incertezas só podem ser eliminados por experimentos mais trabalhosos, e, embora grande parte da sequência tenha sido confirmada, provavelmente ainda esperamos muito tempo antes que possa ser estabelecida a sequência completa. Não estamos certos de que há qualquer justificativa científica para estabelecer cada detalhe e, como achamos que os resultados podem ser úteis para outros pesquisadores, foi decidido publicar a sequência em sua presente forma.¹

E assim eles fizeram. Os dados da sequência foram apresentados em duas páginas que, devido ao modo no qual a publicação foi im-

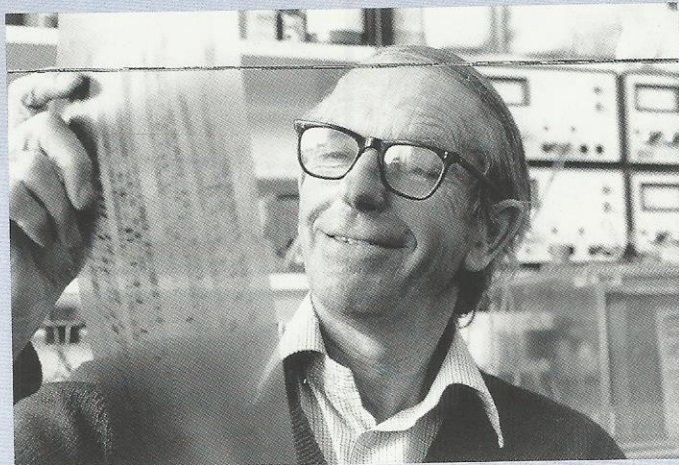


Fig. 1 ■ Frederick Sanger, que desenvolveu técnicas para sequenciamento de moléculas de DNA. Por seu trabalho, Sanger compartilhou o Prêmio Nobel de Química em 1980.

pressa, ficaram bem no meio do trecho citado anteriormente. Assim, os dados foram inseridos em uma desculpa por seu resumo.

A despeito desse resumo, os dados sobre o DNA do Φ X174 foram mais amplos que quaisquer dados de sequência que haviam sido coletados. Eles revelaram a estrutura detalhada de nove genes no genoma do vírus e mostraram que quase todo o DNA do Φ X174 era gênico. Como os autores notaram, “A característica mais marcante do DNA do Φ X174 é o modo em que várias funções do genoma estão comprimidas dentro de 5.375 nucleotídeos.”² De fato, dois dos nove genes foram encontrados residindo dentro de dois outros genes. Assim, no genoma do Φ X174 havia genes dentro de outros genes, um fenômeno que ninguém tinha visto antes. Discutiremos esta descoberta no Cap. 14 quando considerarmos os modos pelos quais os genes podem ser definidos.

Em dezembro de 1977, Sanger e dois de seus colegas relataram uma nova técnica de sequenciamento do DNA.³ Ela é baseada no modo da terminação da síntese do DNA *in vitro* em nucleotídeos específicos – por exemplo, em locais onde uma A deveria ser incorporada à cadeia crescente de DNA ou em locais onde uma T deveria ser incorporada. Analisando os tamanhos das cadeias produzidas nessas condições, Sanger e seus colegas puderam identificar os locais de todas as A, todas as T, e assim por diante. A técnica de término da cadeia foi uma grande melhoria em relação aos métodos anteriores que eram usados para sequenciar o DNA, incluindo o método de mais e menos que havia fornecido a maioria dos dados da sequência do Φ X174. Sanger e seus colegas imediatamente usaram o novo método para esclarecer as incer-

¹Sanger, F., G.M. Air, B.G. Barrell, N.L. Brown, A.R. Coulson, J.C. Fields, C.A. Hutchison III, P.M. Slocumbe, e M. Smith. 1977. Nucleotide sequence of bacteriophage Φ X174 DNA. *Nature* 265:687-695.

²Ibid.

³Sanger, F., S. Nicklen, e A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.



UM MARCO NA GENÉTICA (continuação)

tezas deixadas pelo trabalho anterior. Em 1978, eles publicaram a sequência completa do genoma do Φ X174, um total de 5.386 nucleotídeos.⁴ Nesta época, entretanto, os leitores estavam menos interessados na sequência real do que na técnica que tinha sido usada para obtê-la. A técnica de término de cadeia de Sanger tornou-se o método padrão para sequenciar o DNA. Ainda é usada hoje. No Cap. 15, discutiremos este método de sequenciamento de DNA e no Cap. 16 veremos como ele foi aplicado para sequenciar genomas muito maiores que o genoma do Φ X174. Sem esta técnica, os vários projetos do genoma, incluindo, logicamente, o projeto do genoma humano, nunca teriam sido materializados.

⁴Sanger, F., A.R. Coulson, T. Friedmann, G.M. Air, B.G. Barrell, N.L. Brown, J.C. Fields, C.A. Hutchison III, P.M. Slocombe, e M. Smith. 1978. The nucleotide sequence of bacteriophage Φ X174. *J. Molec. Biol.* 125:225–246.

QUESTÕES PARA DISCUSSÕES

1. Os avanços científicos em geral dependem do desenvolvimento de técnicas-chaves. A técnica de término de cadeia, de Sanger, para sequenciar o DNA é um exemplo. Você pode pensar em outros exemplos?
2. Os vírus podem parecer simples, mas sua organização genética pode ser complexa. O Φ X174 tem dez genes, e não nove como a equipe de Sanger originalmente pensou, e três desses genes estão totalmente inseridos em outros genes. Pelo que você sabe sobre o modo pelo qual a informação genética é codificada e expressa, como um gene pode estar totalmente contido dentro de outro gene? Que desafios o arranjo gene-dentro-de-outro-gene cria para a evolução de um vírus como o Φ X174?

crescimento humano, antes isolado de cadáveres, também é produzido em bactérias. Este hormônio é usado para tratar crianças que não podem produzir quantidades suficientes do hormônio porque possuem um alelo mutante do gene do hormônio de crescimento. Sem a ajuda do hormônio, tais crianças seriam afetadas por nanismo. Muitas outras proteínas de importância médica hoje são rotineiramente produzidas em bactérias que foram transformadas com o gene humano apropriado. A produção em larga escala de tais proteínas é um aspecto da crescente indústria de biotecnologia. Exploraremos modos de produzir proteínas humanas em bactérias no Cap. 15.

A terapia gênica humana é outro modo pelo qual tecnologias genéticas moleculares são usadas para tratar doenças. A estratégia neste tipo de terapia é inserir uma cópia saudável funcional de um determinado gene em células de uma pessoa que só tem cópias mutantes deste gene. O gene inserido pode então compensar os genes defeituosos que a pessoa herdou. Hoje em dia, a terapia gênica humana teve resultados mistos. Os esforços para curar indivíduos com fibrose cística (CF), um grave distúrbio respiratório, introduzindo cópias do gene CF normal em células pulmonares não tiveram sucesso. Entretanto, médicos geneticistas tiveram algum sucesso em tratar distúrbios do sistema imunológico e de células sanguíneas introduzindo os genes normais apropriados em células da medula óssea, que depois se diferenciam em células imunológicas e células sanguíneas. Discutiremos o surgimento de tecnologias para terapia gênica humana e alguns dos riscos envolvidos no Cap. 17.

Genética na Sociedade

As sociedades modernas dependem muito da tecnologia que emerge de pesquisas nas ciências básicas. Nossas indústrias

e serviços estão construindo tecnologias para produção em massa, comunicação instantânea e incríveis processamentos de informação. Nossos estilos de vida também dependem dessas tecnologias. Em um nível mais fundamental, as sociedades modernas estão baseadas na tecnologia para fornecer alimentos e cuidados de saúde. Nós já vimos como a genética está contribuindo para estas importantes necessidades. Entretanto, a genética tem impacto na sociedade também de outras maneiras.

Uma maneira é a econômica. As descobertas das pesquisas genéticas iniciaram inúmeros aspectos comerciais na indústria de biotecnologia. As empresas que fabricam produtos farmacêuticos e testes diagnósticos, ou que fornecem serviços tais como perfis de DNA, têm contribuído para o crescimento econômico mundial. Outra maneira é a legal. As sequências de DNA diferem entre pessoas, e, analisando tais diferenças, as pessoas podem ser identificadas como únicas. Tais análises hoje são rotineiramente usadas em muitas situações – para testes de paternidade, para condenar culpados e inocentar pessoas de crimes dos quais foram acusadas, para autenticar pedidos de heranças e para identificar os mortos. Evidências baseadas na análise do DNA hoje são de uso comum em tribunais de todo o mundo.

Mas o impacto da genética vai além dos aspectos materiais, comerciais e legais de nossas sociedades. Ele atinge o cerne de nossa existência, pois, afinal, o DNA, o tema da genética, é uma parte crucial de nós. As descobertas da genética criam profundas, difíceis e às vezes perturbadoras questões existenciais. Quem somos nós? De onde viemos? A constituição genética determina nossa natureza? nossos talentos? nossa capacidade de aprender? nosso comportamento? Ela tem um papel em estabelecer nossos hábitos? Ela afeta os meios de organização de nossa sociedade? Ela influencia nossas atitudes em relação a outras pessoas? O conheci-

mento de nossos genes e de como eles nos influenciam afeta nossas idéias sobre moralidade e justiça, inocência e culpa, liberdade e responsabilidade? Esse conhecimento muda a maneira como pensamos sobre o que significa ser humano? Gostemos ou não, estas e outras dúvidas nos aguardam em um futuro não muito distante.

PONTOS IMPORTANTES

- As descobertas na genética estão mudando procedimentos e práticas na agricultura e na medicina.
- Os avanços na genética estão criando dúvidas éticas, legais, políticas, sociais e filosóficas.

Exercícios Básicos

Ilustram a análise genética básica.

1. Como a informação genética é expressa nas células?

Resposta: A informação genética é codificada em seqüências no DNA. Inicialmente, essas seqüências são usadas para sintetizar um RNA complementar a elas, um processo chamado transcrição, e então o RNA é usado como molde para especificar a incorporação de aminoácidos na seqüência de um polipeptídeo, um processo chamado tradução. Cada aminoácido no polipeptídeo corresponde a uma seqüência de três nucleotídeos no DNA. As trincas de nucleotídeos que codificam aminoácidos diferentes são chamadas códons.

2. Qual o significado evolutivo de uma mutação?

Resposta: A mutação cria uma variação nas seqüências de DNA dos genes (e também nos componentes não genéticos dos genomas). Essa variação acumula-se nas populações de organismos com o tempo e pode acabar por produzir diferenças observáveis entre os organismos. Uma população pode vir a diferir de outra de acordo com os tipos de mutações que se acumularam com o tempo. Assim, a mutação fornece o impulso para resultados evolutivos diferentes na população.

Testar Seus Conhecimentos

Integra conceitos e técnicas diferentes.

1. Suponha que um gene contém 10 códons. Quantos nucleotídeos codificantes o gene contém? Quantos aminoácidos se espera estar presentes nesse produto polipeptídico? Entre todos os possíveis genes compostos de 10 códons, quantos polipeptídeos diferentes poderiam ser produzidos?

Resposta: O gene possui 30 nucleotídeos codificantes. Espera-se que seu produto polipeptídico contenha 10 aminoá-

cidos, cada um correspondendo a um dos códons no gene. Se cada códon pode especificar um de 20 aminoácidos de ocorrência natural, entre todas as possíveis seqüências de genes com 10 códons de tamanho, podemos imaginar um total de 20^{10} produtos polipeptídicos, um número verdadeiramente enorme!

Questões e Problemas

Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 1.1. Em poucas frases, quais eram as idéias principais de Mendel sobre a herança?
- 1.2. Tanto o DNA quanto o RNA são compostos de nucleotídeos. Que moléculas se combinam para formar um nucleotídeo?
- 1.3. Que bases estão presentes no DNA? Que bases estão presentes no RNA? Que açúcares estão presentes em cada um destes ácidos nucleicos?
- 1.4. O que é um genoma?
- 1.5. A seqüência de um filamento de DNA é ATTGCCGTC. Se este filamento servir como molde para a síntese de DNA, qual será a seqüência do filamento recém-sintetizado?
- 1.6. Um gene contém 141 códons. Quantos nucleotídeos estão presentes na seqüência codificante do gene? Quantos aminoácidos se espera estar presentes no polipeptídeo codificado por este gene?
- 1.7. O filamento molde de um gene que está sendo transcrito é CTTGCCAGT. Qual será a seqüência do RNA feito a partir deste molde?
- 1.8. Qual a diferença entre transcrição e tradução?
- 1.9. O RNA é sintetizado usando um molde de DNA. O DNA alguma vez é sintetizado usando o RNA como molde? Explique.
- 1.10. A anemia falciforme é causada por uma mutação em um dos códons no gene para β -globina; devido a esta mutação, o

sexto aminoácido no polipeptídeo de β -globina é uma valina em vez de um ácido glutâmico. Um tipo menos grave de anemia é causado por uma mutação que troca este mesmo códon por um que especifica lisina como o sexto aminoácido no polipeptídeo de β -globina. Que palavra é usada para descrever as duas formas mutantes deste gene? Você acha que uma pessoa que tenha estas duas formas mutantes do gene de β -globina sofreria anemia? Explique.

- 1.11.** O gene para α -globina está presente em todas as espécies de vertebrados. Durante milhões de anos, a sequência de DNA deste gene mudou na linhagem de cada espécie. Consequentemente, a sequência de aminoácidos de α -globina também mudou nessas linhagens. Entre as 141 posições de aminoácidos neste polipeptídeo, a α -globina humana difere da de tubarão em 79 posições; difere da

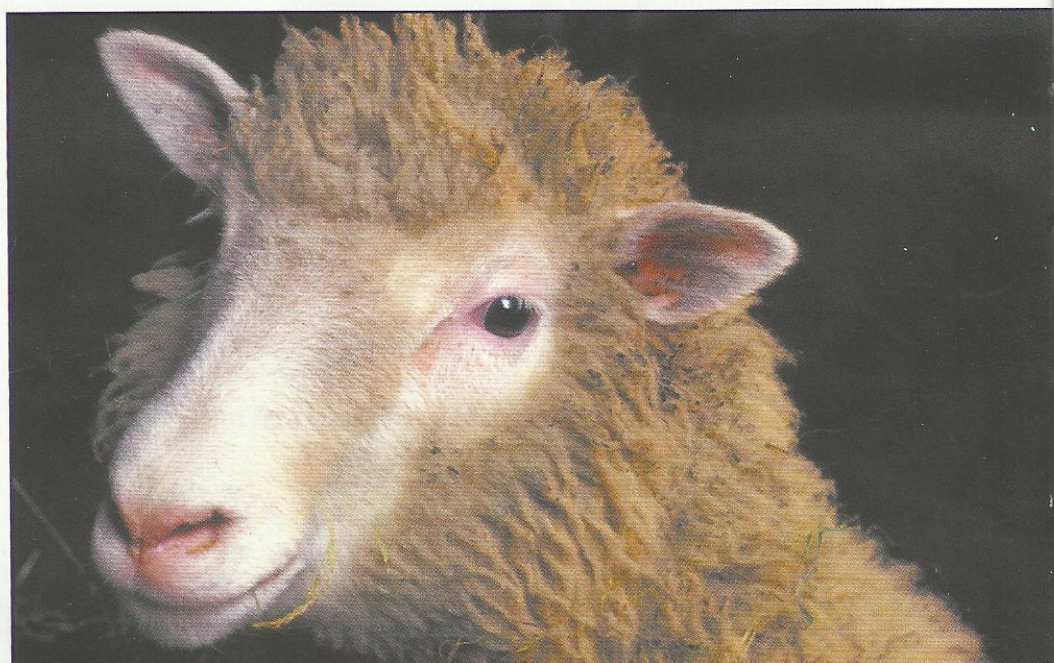
de carpa em 68; e difere da de vaca em 17. Estes dados sugerem uma filogenia evolutiva para tais espécies de vertebrados?

- 1.12.** A hemofilia é um distúrbio hereditário no qual o mecanismo de coagulação do sangue é defeituoso. Devido a este defeito, pessoas com hemofilia podem morrer de cortes ou contusões, especialmente se forem lesados órgãos internos tais como fígado, pulmões ou rins. Um método para tratamento envolve injetar um fator de coagulação que foi purificado de doadores de sangue. Esse fator é uma proteína codificada por um gene humano. Sugira um modo pelo qual a moderna tecnologia genética poderia ser usada para produzir este fator em escala industrial. Existe um modo pelo qual o erro inato da hemofilia possa ser corrigido pela terapia gênica humana?

Reprodução Celular e Organismos Genéticos Modelos

Tópicos

- Células e Cromossomos
- Mitose
- Meiose
- Genética no Laboratório: Uma Introdução a Alguns Organismos Modelos de Pesquisas



Dolly, o primeiro mamífero clonado.

Dolly

As ovelhas vivem nos campos da Escócia há séculos. As Finn Dorsets e as Scottish Blackfaces são algumas das raças criadas lá pelos pastores. A cada primavera, as ovelhas que foram concebidas durante o outono nascem. Elas crescem rapidamente e vivem em rebanhos ou em criadouros. No início de 1997, uma ovelha diferente de qualquer outra veio ao mundo. Essa ovelha, chamada Dolly, não teve um pai, mas ela teve três mães; além disso, seus genes eram idênticos aos de uma de suas mães. Em uma palavra, Dolly era um clone.

Cientistas do Roslin Institute perto de Edimburgo, Escócia, produziram Dolly ao fundir um ovócito de uma fêmea Blackface (a célula mãe do ovócito) com uma célula da úbere de uma fêmea Finn Dorset (a mãe genética). O material genético do ovócito da fêmea Blackface tinha sido removido antes da fusão do ovócito com a célula da úbere. Subseqüentemente, o novo "zigoto" foi estimulado a se dividir. Ele produziu um embrião, que foi implantado no útero

de outra fêmea Blackface (a mãe gestacional ou adotiva). Esse embrião cresceu e se desenvolveu, e, quando a gestação da mãe adotiva chegou a termo, nasceu Dolly.

A tecnologia que produziu Dolly surgiu de um século de pesquisas básicas sobre a base celular da reprodução. No curso comum de eventos, um ovócito de uma fêmea é fertilizado por um espermatozóide de um macho, e o zigoto resultante se divide para produzir células geneticamente idênticas. Tais células então se dividem muitas vezes para produzir um organismo multicelular. Dentro desse organismo, um grupo particular de células segue um modo diferente de divisão para produzir células reprodutivas especializadas – ovócitos ou espermatozóides. Um ovócito de tal organismo então se une a um espermatozóide de outro organismo para produzir uma nova prole. A prole cresce e o ciclo continua, geração após geração. Mas, entre as ovelhas dos pastos da Escócia, Dolly não seguiu este padrão de reprodução. Claramente, seus criadores tinham algo em mente.

CÉLULAS E CROMOSSOMOS

Tanto nas células procarióticas quanto nas eucarióticas, o material genético é organizado em cromossomos.



Na primeira parte do século dezenove, algumas décadas antes que Gregor Mendel fizesse seus experimentos com ervilhas, biólogos estabeleceram o princípio de que os seres vivos são compostos de células. Alguns organismos consistem em apenas uma célula. Outros consistem em trilhões de células. Cada célula é uma reunião complicada de moléculas que podem adquirir materiais, obter e estocar energia e desempenhar atividades diversas, incluindo reprodução. As formas mais simples de vida, os vírus, não são compostos de células. Entretanto, os vírus devem penetrar em células para poder desempenhar sua função. Assim, toda vida tem uma base celular. Como preparação para nossa jornada através da ciência da genética, agora vamos rever a biologia das células. Também discutiremos os cromossomos – as estruturas celulares nas quais residem os genes.

O Ambiente Celular

As células vivas são feitas de muitos tipos diferentes de moléculas. A mais abundante é a água. Moléculas pequenas, como sais, açúcares, aminoácidos e algumas vitaminas, prontamente se dissolvem em água, e algumas moléculas maiores interagem favoravelmente com ela. Todos estes tipos de substâncias são ditas hidrofílicas. Outros tipos de moléculas não interagem bem com a água. Elas são ditas hidrofóbicas. O interior de uma célula, chamado **citoplasma**, contém tanto substâncias hidrofílicas quanto hidrofóbicas.

As moléculas que constituem as células são diferentes em estrutura e função.

Os **carboidratos** tais como amido e glicogênio estocam energia química para trabalho dentro das células. Essas moléculas são compostas de glicose, um açúcar simples. As subunidades de glicose são ligadas umas às outras para formar cadeias longas, ou polímeros. As células obtêm energia quando as moléculas de glicose liberadas dessas cadeias são quimicamente degradadas em compostos mais simples e, finalmente, em dióxido de carbono e água. As células também possuem uma variedade de **lipídeos**. Tais moléculas são formadas por interações químicas entre glicerol, um pequeno composto orgânico, e compostos orgânicos maiores chamados ácidos graxos. Os lipídeos são constituintes importantes de muitas estruturas dentro das células. Eles também servem como fontes de energia. **Proteínas** são as mais diversas moléculas dentro das células. Cada proteína consiste em um ou mais polipeptídeos, que são cadeias de aminoácidos. Geralmente uma proteína consiste em dois polipeptídeos, e esta proteína é um dímero; às vezes, a proteína consiste em muitos polipeptídeos, isto é, é um multímero. Dentro das células, as proteínas são componentes de muitas estruturas diferentes. Elas também catalisam reações químicas. Chamamos tais proteínas catalíticas de **enzimas**. As células também contêm ácidos nucleicos, DNA e RNA, os quais, como descrito no Cap. 1, são essenciais para a vida.

As células são circundadas por uma fina camada chamada **membrana**. Muitos tipos diferentes de moléculas constituem as membranas celulares; entretanto, os constituintes primários são lipídeos e proteínas. Membranas também estão presentes dentro das células. Essas membranas internas podem dividir a célula em compartimentos ou podem ajudar a formar estruturas especializadas chamadas **organelas**. As membranas são fluidas e flexíveis. Muitas das moléculas dentro de uma membrana não são rigidamente mantidas no lugar por forças químicas fortes. Consequentemente, elas são capazes de deslizar umas sobre as outras no que seria um mar molecular em constante mudança. Alguns tipos de células são circundadas por paredes rígidas, que são **externas** à

membrana. As paredes das células vegetais são compostas de celulose, um carboidrato complexo. As paredes das bactérias são compostas de um tipo diferente de material chamado mureína.

As paredes e as membranas separam os conteúdos de uma célula do mundo externo. Entretanto, elas não a vedam. Tais estruturas são porosas a alguns materiais e permitem seletivamente que outros materiais passem através delas por canais e comportas. O transporte de materiais nas paredes e membranas e através delas é uma atividade importante das células. As membranas celulares também contêm moléculas que interagem com materiais no ambiente externo da célula. Tais moléculas dão às células informações vitais sobre condições no ambiente e também medeiam atividades celulares importantes.

Células Procarióticas e Eucarióticas

Quando examinamos o mundo vivo, encontramos dois tipos básicos de células: procarióticas e eucarióticas (Fig. 2.1). As **células procarióticas** em geral têm menos de um milésimo de milímetro de tamanho e tipicamente não têm um sistema complicado de membranas internas e organelas membranosas. Seu material hereditário, isto é, o DNA, não é isolado em um compartimento subcelular especial. Os organismos com este tipo de organização celular são chamados procariontes. Exemplos incluem as bactérias, que são as formas de vida mais abundantes no planeta, e as archaea, que são encontradas em ambientes extremos tais como lagos salgados, águas quentes e material vulcânico de alto mar. Todos os outros organismos – plantas, animais, protistas e fungos – são eucariontes.

As **células eucarióticas** (Fig. 2.2) são maiores que as células procarióticas, em geral pelo menos 10 vezes maiores,

e possuem sistemas complicados de membranas internas, algumas das quais estão associadas a organelas conspícuas. Por exemplo, as células eucarióticas contêm tipicamente uma ou mais **mitocôndrias**, que são organelas elipsóides dedicadas à obtenção de energia dos alimentos. As células de algas e plantas contêm um outro tipo de organela para obtenção de energia chamada **cloroplasto**, que capta energia solar e a converte em energia química. Tanto as mitocôndrias quanto os cloroplastos são circundados por membranas.

Um marco de todas as células eucarióticas é que seu material hereditário está contido dentro de uma estrutura delimitada por membrana chamada **núcleo**. Dentro do núcleo, o DNA é organizado em estruturas distintas chamadas **cromossomos**. Os cromossomos individuais podem ser visíveis, dependendo de quando a célula é examinada. Nas células procarióticas, o DNA geralmente não é abrigado dentro de um núcleo bem definido. Investigaremos os modos pelos quais o DNA cromossômico é organizado em células procarióticas e eucarióticas no Cap. 9.

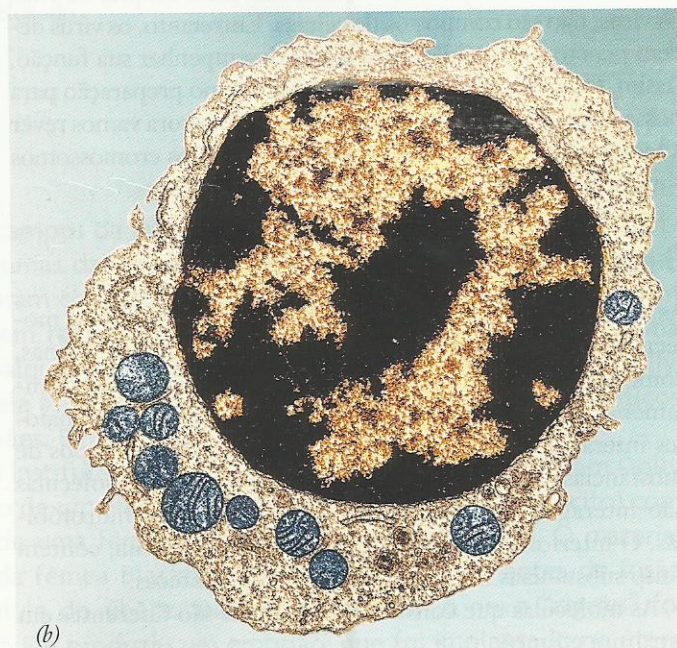
Parte do DNA dentro de uma célula eucariótica não está situado dentro do núcleo. Esse DNA extranuclear está situado nas mitocôndrias e nos cloroplastos. Examinaremos suas estrutura e funcionamento no Cap. 19.

Tanto as células procarióticas quanto as eucarióticas possuem numerosos **ribossomos**, que são organelas pequenas envolvidas na síntese de proteínas, um processo que estudaremos no Cap. 12. Os ribossomos são encontrados no citoplasma. Embora os ribossomos não sejam compostos de membranas, em células eucarióticas eles estão geralmente associados a um sistema de membranas chamado **retículo endoplasmático**. O retículo pode ser conectado ao **complexo de Golgi**, um conjunto de sacos membranosos e vesículas que estão envolvidos na modificação química e no transporte de substâncias dentro das células. Outras pe-



(a)

Fig. 2.1 ■ (a) Eletromicrografia da bactéria *Escherichia coli*, um procarionte, dividindo-se em duas células. O material claro dentro das células é onde está situado o DNA. (b) Eletromicrografia de um leucócito humano mostrando o núcleo (grande área escura), que é característico de células eucarióticas. O material de cor clara dentro do núcleo constitui os cromossomos, que não podem ser vistos individualmente pois não se condensaram na preparação para a divisão celular. Os pequenos círculos azuis são mitocôndrias, e as estruturas finas que parecem atravessá-las são membranas mitocondriais.



(b)

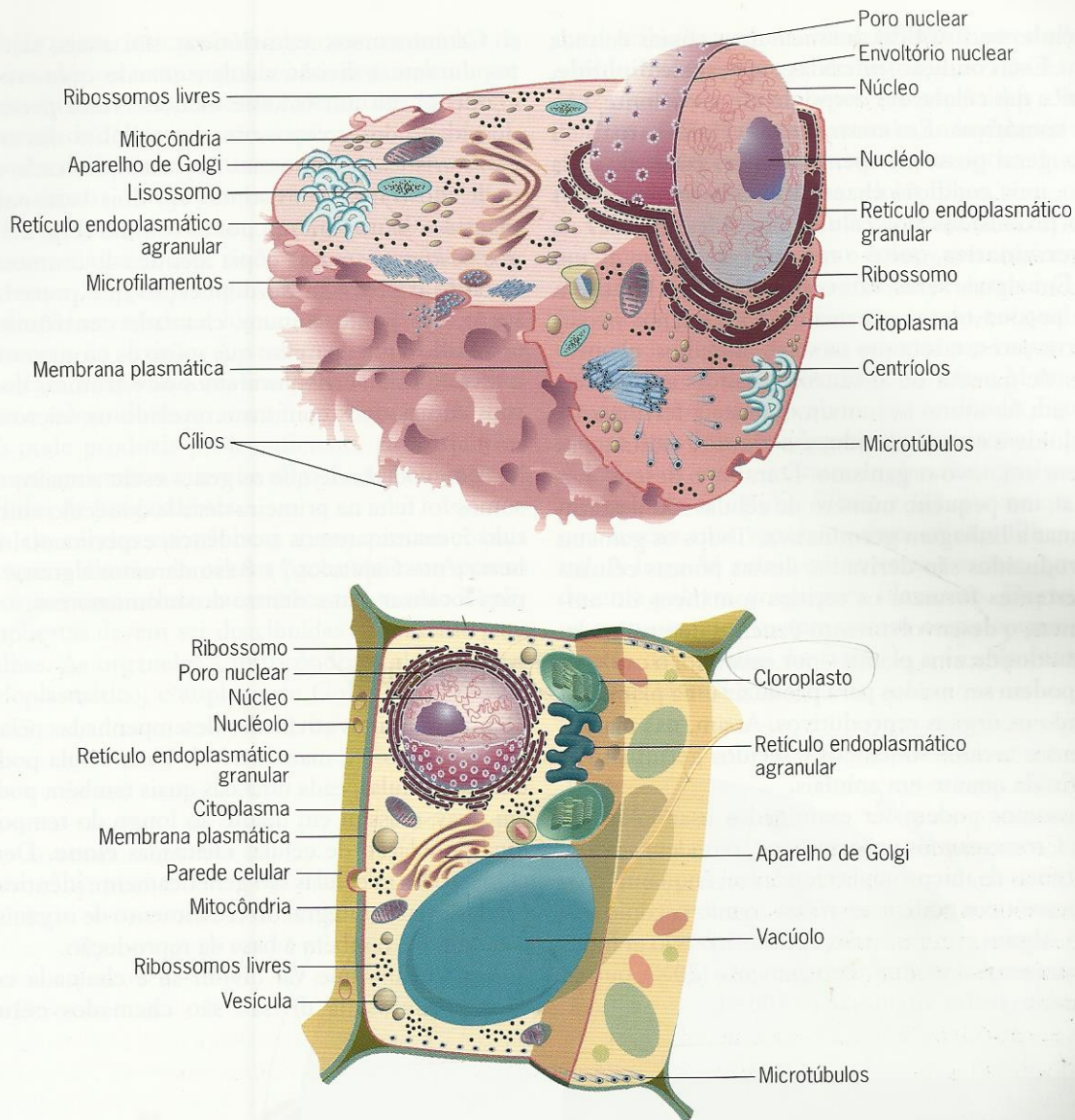


Fig. 2.2 ■ As estruturas de células vegetais e animais.

pequenas organelas delimitadas por membrana podem também ser encontradas em células eucarióticas. Em células animais, os **lisossomos** são produzidos pelo complexo de Golgi. Tais organelas contêm tipos diferentes de enzimas digestivas que destruiriam a célula se fossem liberadas no citoplasma. Tanto células vegetais quanto animais podem conter **peroxissomos**, que são pequenas organelas dedicadas ao metabolismo de substâncias tais como gorduras e aminoácidos.

As formas e atividades das células eucarióticas são influenciadas por um sistema de filamentos, fibras e moléculas associadas que coletivamente formam o **citoesqueleto**. Esses materiais dão forma às células e possibilitam que alguns tipos de células se movam em seu ambiente, um fenômeno chamado **motilidade celular**. O citoesqueleto mantém as organelas no lugar e tem um papel importante em mover materiais para locais específicos dentro das células, um fenômeno chamado **tráfego**.

Cromossomos: Onde Estão Situados os Genes

Cada cromossomo consiste em uma molécula de dupla hélice de DNA mais um conjunto de proteínas. O RNA pode também estar associado a cromossomos. As células procarióticas contêm tipicamente apenas um cromossomo, embora algumas vezes possuam muitas moléculas menores de DNA chamadas **plasmídeos**. A maioria das células eucarióticas contêm vários cromossomos diferentes – por exemplo, os espermatozoides humanos têm 23. Os cromossomos de células eucarióticas também são tipicamente maiores e mais complexos que os das células procarióticas. As moléculas de DNA em cromossomos procarióticos e plasmídeos são circulares, como são a maioria das moléculas de DNA encontradas em mitocôndrias e cloroplastos de células eucarióticas. Entretanto, as moléculas de DNA encontradas nos cromossomos nos núcleos de células eucarióticas são lineares.

Muitas células eucarióticas possuem duas cópias de cada cromossomo. Esta condição, referida como estado **diplóide**, é característica das células no corpo de um eucarionte, isto é, as células **somáticas**. Em contraste, as células sexuais ou **gametas** em geral possuem apenas uma só cópia de cada cromossomo, uma condição chamada estado **haplóide**. Os gametas são produzidos por células diplóides situadas na **linhagem germinativa**, que é o tecido reprodutivo de um organismo. Em alguns seres, tais como plantas, a linhagem germinativa produz tanto espermatozóides quanto ovócitos. Em outros seres, tais como os seres humanos, ela produz um tipo de gameta ou o outro. Quando um gameta masculino e um feminino se juntam durante a fertilização, o estado diplóide é reestabelecido, e o zigoto resultante se desenvolve em um novo organismo. Durante o desenvolvimento animal, um pequeno número de células são destacadas para formar a linhagem germinativa. Todos os gametas que serão produzidos são derivados dessas poucas células. As células restantes formam os tecidos somáticos do animal. Em plantas, o desenvolvimento é menos determinado. Os tecidos obtidos de uma planta – por exemplo, o caule ou uma folha – podem ser usados para produzir uma planta inteira, incluindo os órgãos reprodutivos. Assim, nas plantas a distinção entre tecidos somáticos e tecidos germinativos não é tão definida quanto em animais.

Os cromossomos podem ser examinados usando-se um microscópio. Cromossomos procarióticos só podem ser vistos com as técnicas de microscopia eletrônica, enquanto cromossomos eucarióticos podem ser vistos ao microscópio óptico (Fig. 2.3). Alguns cromossomos eucarióticos são grandes o suficiente para ser vistos com baixo aumento (20×); outros precisam de muito maior ampliação (>500×).



Cromossomos eucarióticos são mais claramente vistos durante a divisão celular, quando cada cromossomo se condensa em um volume menor. Nesse período, a maior densidade dos cromossomos possibilita discernir algumas características estruturais. Por exemplo, cada cromossomo pode mostrar-se consistindo em dois bastonetes paralelos mantidos juntos em um ponto comum (Fig. 2.3b). Cada um dos bastonetes é uma cópia idêntica do cromossomo criado durante o processo de duplicação que precede a condensação, e o ponto comum, chamado **centrômero**, torna-se associado a um aparelho que move os cromossomos durante a divisão celular. Exploraremos as estruturas dos cromossomos eucarióticos conforme revelado na microscopia óptica no Cap. 6.

A descoberta de que os genes estão situados nos cromossomos foi feita na primeira década do século vinte. No Capítulo 5, examinaremos a evidência experimental desta descoberta e nos Capítulos 7 e 8 estudaremos algumas das técnicas para localizar genes dentro dos cromossomos.

Divisão Celular

Dentre as muitas atividades desempenhadas pelas células vivas, a divisão é a mais incrível. Uma célula pode dividir-se em duas células, cada uma das quais também pode dividir-se em duas, e assim em diante ao longo do tempo, para criar uma população de células chamadas **clone**. Dentro de um clone, todas as células são geneticamente idênticas. A divisão celular é parte integral do crescimento de organismos multicelulares e é também a base da reprodução.

Uma célula que vai dividir-se é chamada **célula mãe**, e os produtos da divisão são chamados **células filhas**.



Fig. 2.3 ■ (a) Eletromicrografia mostrando um cromossomo bacteriano. (b) Micrografia óptica de cromossomos humanos durante divisão celular.

Quando células procarióticas se dividem, os conteúdos da célula mãe são mais ou menos igualmente repartidos entre as duas células filhas. Este processo é denominado **fissão**. Os cromossomos da célula mãe são duplicados antes da divisão, e as cópias deles são repartidas entre as células filhas. Em condições ótimas, um procarionte tal como a bactéria intestinal *Escherichia coli* divide-se a cada 20 ou 30 minutos. Nesta taxa, uma única *E. coli* pode formar um clone de aproximadamente 2^{50} células, mais de um quatrilhão, em apenas um dia. Na verdade, é lógico, as *E. coli* não mantêm esta alta taxa de divisão. À medida que as células se acumulam, a taxa de divisão declina porque os nutrientes são exauridos e subprodutos se acumulam. Entretanto, uma única *E. coli* pode produzir prole suficiente em um único dia para formar uma massa visível a olho nu. Chamamos tal massa de células de uma **colônia**.

A divisão de células eucarióticas é um processo mais complicado que a divisão de células procarióticas. Após os cromossomos em uma célula eucariótica terem sido duplicados, as duplicatas devem ser distribuídas igualmente para as células filhas. As organelas – mitocôndrias, cloroplastos, retículo endoplasmático, complexo de Golgi etc. – devem

ser distribuídas para as células filhas. Entretanto, para estes constituintes o processo de distribuição não é igual e exato. Mitocôndrias e cloroplastos são repartidos aleatoriamente para as células filhas. O retículo endoplasmático e o complexo de Golgi são fragmentados na época da divisão e depois são reconstituídos nas células filhas.

Cada vez que uma célula eucariótica se divide, ela passa por uma série de fases que coletivamente formam o **ciclo celular** (Fig. 2.4). A progressão de fases é $G_1 \rightarrow S \rightarrow G_2 \rightarrow M$. Nesta progressão, S é o período no qual os cromossomos são duplicados, um evento que requer a síntese de DNA, representado pela letra S. A fase M no ciclo celular é a ocasião em que a célula mãe se divide. Esta fase geralmente tem dois componentes: (1) **mitose**, o processo que distribui os cromossomos duplicados igualmente e exatamente para as células filhas, e (2) **citocinese**, o processo que fisicamente separa as duas células filhas uma da outra. A letra M refere-se ao termo *mitose*, que é derivado de uma palavra grega para filamento. Durante a mitose, os cromossomos se apresentam como corpos filamentosos dentro das células. As fases G_1 e G_2 são intervalos (*gaps*) entre as fases S e M.

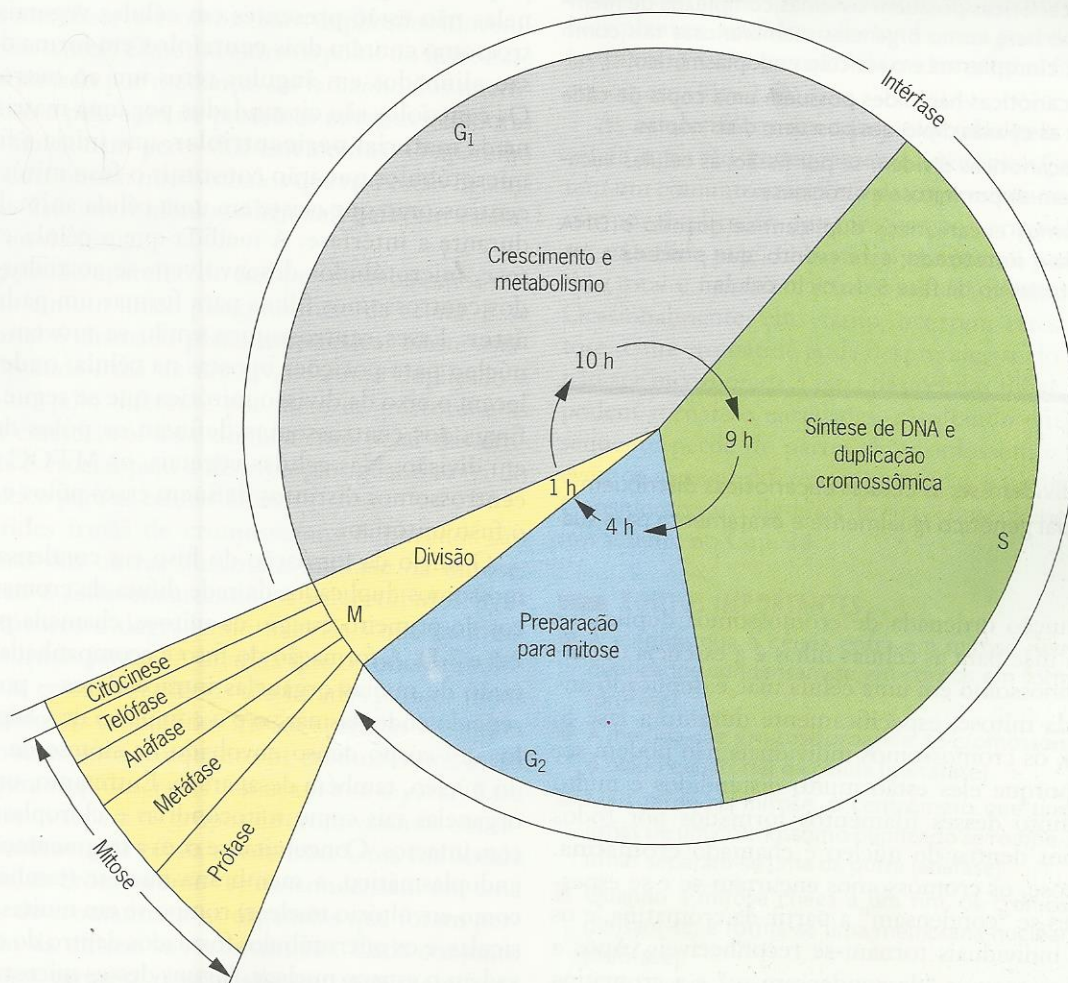


Fig. 2.4 ■ O ciclo celular eucariótico. Este ciclo tem 24 horas de duração. Essa duração varia entre tipos diferentes de células eucarióticas.

A duração de um ciclo celular varia entre os diferentes tipos de células. Em embriões, nos quais o crescimento é rápido, o ciclo pode ser tão curto quanto 30 minutos. Em tecidos adultos de crescimento lento, ele pode durar vários meses. Algumas células, tais como as do tecido nervoso e muscular, param de se dividir quando adquirem suas funções especializadas. A progressão das células eucarióticas pelo ciclo é rigidamente controlada por tipos diferentes de proteínas. Quando as atividades dessas proteínas são perturbadas, as células dividem-se de modo irregular. Essa desregulação da divisão celular pode resultar em câncer, que é uma importante causa de morte entre as pessoas hoje em dia. No Cap. 24, investigaremos a base genética do câncer.

PONTOS IMPORTANTES

- As células, as unidades básicas de todos os seres vivos, são circundadas por membranas.
- Os cromossomos, as estruturas celulares que levam os genes, são compostos de DNA, RNA e proteína.
- Em eucariontes, os cromossomos estão contidos dentro de um núcleo delimitado por membrana, o que não ocorrem em procariontes.
- As células eucarióticas possuem sistemas complexos de membranas internas bem como organelas membranosas tais como mitocôndrias, cloroplastos e o retículo endoplasmático.
- As células eucarióticas haplóides possuem uma cópia de cada cromossomo; as células diplóides possuem duas cópias.
- As células procarióticas dividem-se por fissão; as células eucarióticas dividem-se por mitose e citocinese.
- Os cromossomos eucarióticos duplicam-se quando o DNA de uma célula é sintetizado; este evento, que precede a mitose, é característico da fase S do ciclo celular.

MITOSE

Quando dividem-se, as células eucarióticas distribuem seu material genético igualmente e exatamente para sua prole.

A distribuição ordenada de cromossomos duplicados em uma célula mãe para as células filhas é a essência da mitose. Cada cromossomo em uma célula mãe é duplicado antes do início da mitose, especificamente durante a fase S. Nesse período, os cromossomos individuais não podem ser identificados porque eles estão muito distendidos e muito finos. O conjunto desses filamentos formados por todos os cromossomos dentro do núcleo é chamado **cromatina**. Durante a mitose, os cromossomos encurtam-se e se espessam, isto é, eles se “condensam” a partir da cromatina, e os cromossomos individuais tornam-se reconhecíveis. Após a mitose, os cromossomos “descondensam-se” e a cromatina é reconstituída. Os biólogos referem-se ao período no qual os cromossomos individuais não podem ser vistos como a

intérfase. Este período, que pode ser muito longo, é a época entre os sucessivos eventos mitóticos.

Quando a mitose começa, cada cromossomo já está duplicado. As duplicatas, chamadas **cromátides irmãs**, permanecem intimamente associadas umas às outras e estão unidas pelo centrômero do cromossomo. O termo *irmãs* é de certo modo incorreto, pois tais cromátides são cópias exatas do cromossomo original, e portanto são mais idênticas do que irmãs. Talvez a palavra “gêmeas” descreveria melhor a situação. Entretanto, o termo “irmãs” é comumente usado, e o manteremos aqui.

A distribuição de cromossomos duplicados para as células filhas é organizada e executada pelos **microtúbulos**, que são componentes do citoesqueleto. Essas fibras, compostas de proteínas chamadas tubulinas, ligam-se aos cromossomos e os movimentam durante a divisão da célula mãe. Durante a mitose, os microtúbulos reúnem-se em um arranjo complexo chamado **fuso** (Fig. 2.5a). A formação do fuso depende dos **centros de organização de microtúbulos (MTOC)**, que são encontrados no citoplasma de células eucarióticas, geralmente perto do núcleo. Nas células animais, os MTOC são diferenciados em pequenas organelas chamadas **centrossomos**. Tais organelas não estão presentes em células vegetais. Cada centrossomo contém dois **centríolos** em forma de barril, que são alinhados em ângulos retos um ao outro (Fig. 2.5b). Os centríolos são circundados por uma matriz difusa chamada **material pericentriolar**, que inicia a formação dos microtúbulos que irão constituir o fuso mitótico. O único centrossomo que existe em uma célula animal é duplicado durante a intérfase. À medida que a célula entra em mitose, microtúbulos desenvolvem-se ao redor de cada um dos centrossomos filhos para formar um padrão chamado **âster**. Esses centrossomos então se movem ao redor do núcleo para posições opostas na célula, onde eles estabelecem o eixo da divisão mitótica que se segue. As posições finais dos centrossomos definem os pólos da célula mãe em divisão. Nas células vegetais, os MTOC que não têm centrossomos distintos definem esses pólos e estabelecem o fuso mitótico.

O início da formação do fuso e a condensação dos cromossomos duplicados da rede difusa de cromatina são marcos do primeiro estágio da mitose, chamada **prófase** (Figs. 2.6 e 2.7). A formação do fuso é acompanhada de fragmentação de muitas organelas intracelulares – por exemplo, o retículo endoplasmático e o complexo de Golgi. O **núcleo**, um corpo denso envolvido na síntese de RNA dentro do núcleo, também desaparece. Entretanto, outros tipos de organelas tais como mitocôndrias e cloroplastos permanecem intactos. Concomitante com a fragmentação do retículo endoplasmático, a membrana nuclear (também conhecida como envoltório nuclear) rompe-se em muitas pequenas vesículas, e os microtúbulos formados dentro do citoplasma invadem o espaço nuclear. Alguns desses microtúbulos ligam-se aos **cinetócoros**, que são estruturas protéicas associadas aos centrômeros dos cromossomos duplicados. A ligação de

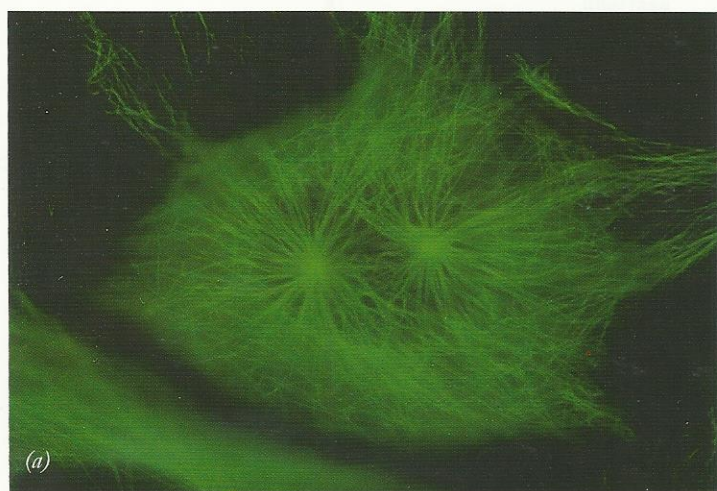


Fig. 2.5 ■ (a) O fuso mitótico em uma célula animal cultivada, que foi corado para mostrar os microtúbulos (verde) saindo dos dois ásteres. (b) Eletromicrografia mostrando dois pares de centríolos.

microtúbulos do fuso aos cinetócoros indica que a célula está entrando na **metáfase** da mitose.

Durante a metáfase, os cromossomos duplicados movem-se para posições a meio caminho entre os pólos do fuso. Esse movimento é ativado por mudanças no tamanho dos microtúbulos do fuso e pela ação de proteínas motoras geradoras de força que funcionam perto dos cinetócoros. O aparelho do fuso também contém microtúbulos que não estão ligados a cinetócoros. Tais microtúbulos adicionais parecem estabilizar o aparelho do fuso.

Pela operação do aparelho do fuso, os cromossomos duplicados ficam em um único plano no meio da célula. Esse plano equatorial é chamado **placa metafásica**. Neste estágio, cada cromátide irmã de um cromossomo duplicado é conectada a um pólo diferente por meio de microtúbulos ligados a seu cinetócoro. Esse alinhamento polar das cromátides irmãs é crucial para a distribuição igual e exata do material genético para as células filhas.

As cromátides irmãs de cromossomos duplicados são separadas umas das outras durante a **anáfase** da mitose. Essa separação é feita encurtando os microtúbulos ligados aos cinetócoros e degradando materiais que mantêm as cromátides irmãs juntas. À medida que os microtúbulos se encurtam, as cromátides irmãs são levadas para os pólos opostos da célula. As cromátides irmãs separadas são agora chamadas cromossomos. Enquanto os cromossomos estão se movendo para os pólos, estes começam a se separar. Esse duplo movimento separa os grupos de cromossomos em espaços distintos dentro da célula em divisão. Uma vez obtida essa separação, os cromossomos se descondensam em uma rede de fibras de cromatina, e as organelas que foram perdidas no início da mitose se reconstituem. Cada conjunto de cromossomos torna-se circundado por uma membrana nuclear. A descondensação dos cromossomos e a restauração das organelas internas são características da **telófase**

da mitose. Quando a mitose está completa, as duas células filhas separam-se pela formação de membranas entre elas. Nas plantas, também é formada uma parede entre as células filhas. Essa separação física das células filhas é chamada citocinese (Fig. 2.8).

As células filhas que são produzidas pela divisão de uma célula mãe são geneticamente idênticas. Cada célula filha tem um conjunto completo de cromossomos que são derivados por duplicação de cromossomos originalmente presentes na célula mãe. O material genético é portanto transmitido total e fielmente para as células filhas pela célula mãe. Ocasionalmente, entretanto, ocorrem erros durante a mitose. Uma cromátide pode desprender-se do fuso mitótico e não ser incorporada a uma das células filhas, ou cromátides podem tornar-se enredadas, resultando em quebra e subsequente perda de partes do cromossomo. Esses tipos de eventos causam diferenças genéticas entre as células filhas. Consideraremos algumas de suas consequências no Cap. 6 e novamente no Cap. 24.

PONTOS IMPORTANTES

- À medida que a célula entra em mitose, seus cromossomos duplicados condensam-se em corpos em forma de bastonete (prófase).
- À medida que a mitose progride, os cromossomos migram para o plano equatorial da célula (metáfase).
- Mais tarde na mitose, o centrômero que une as cromátides irmãs de um cromossomo duplicado se rompe, e as cromátides irmãs separam-se uma da outra (anáfase).
- Quando a mitose chega a um fim, os cromossomos descondensam-se, e forma-se uma membrana nuclear ao redor deles (telófase).
- Cada célula filha produzida por mitose e citocinese tem o mesmo conjunto de cromossomos; assim, as células filhas são geneticamente idênticas.

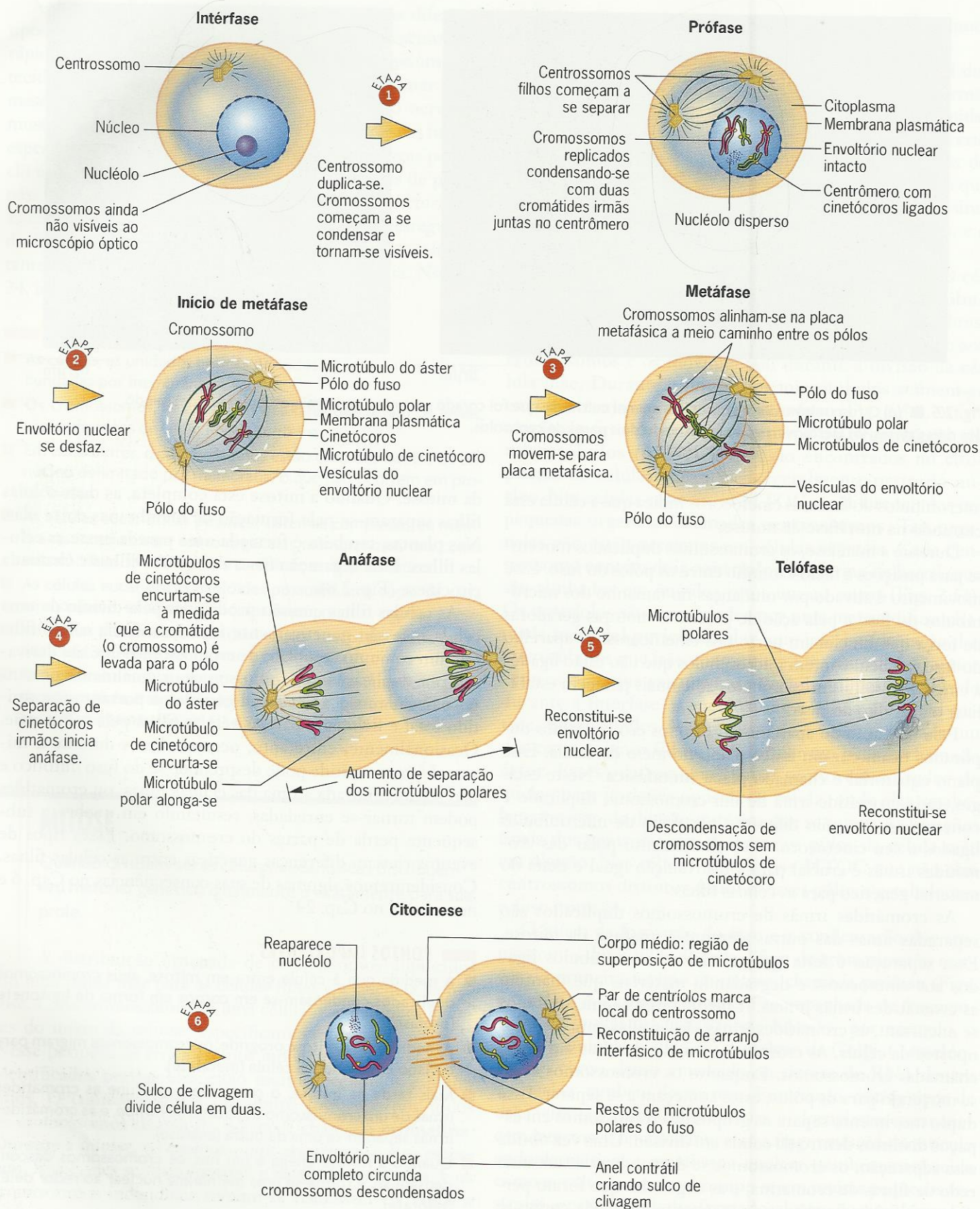


Fig. 2.6 ■ Mitose em células animais.

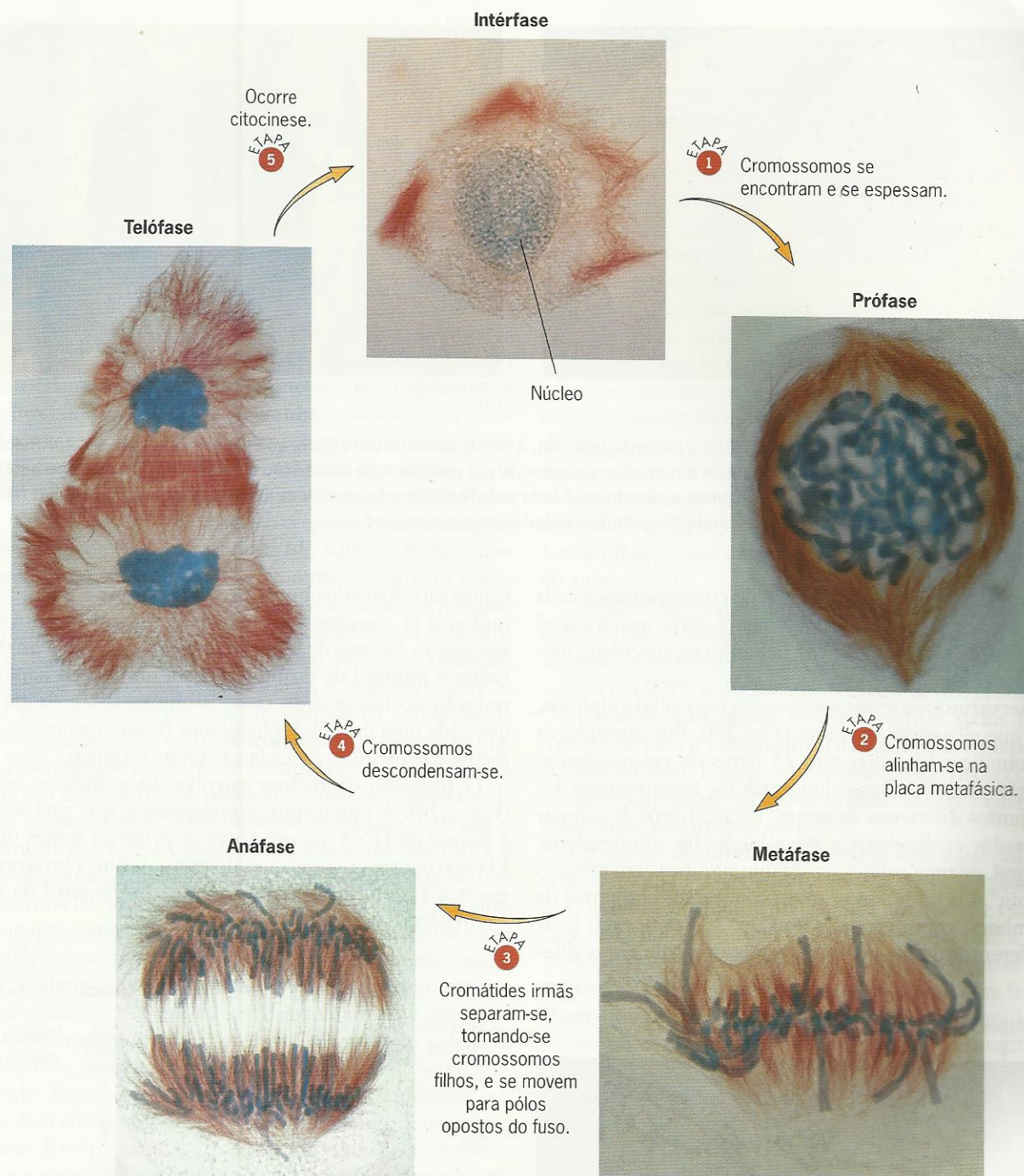


Fig. 2.7 ■ Fotografias dos principais estágios de mitose no lírio *Haemanthus*.

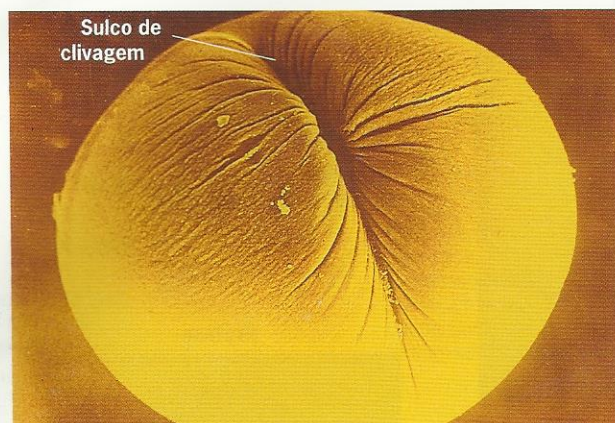
MEIOSE

A reprodução sexual envolve um mecanismo que reduz o número de cromossomos à metade.

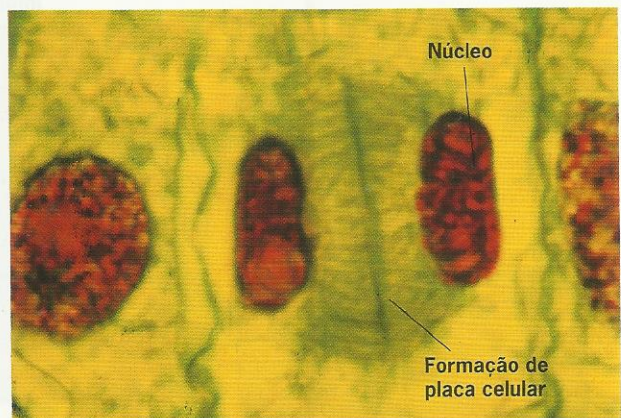


Se indicarmos o número de cromossomos em um gameta pela letra n , então o zigoto produzido pela união dos dois gametas tem $2n$ cromossomos. Cha-

mamos os n cromossomos de um gameta como a condição haplóide e os $2n$ cromossomos do zigoto como o estado diplóide. **Meiose**, da palavra grega que significa “diminuição”, é o processo que reduz o estado diplóide ao estado haplóide, isto é, ele reduz o número de cromossomos de uma célula pela metade. As células haplóides resultantes ou se tornam diretamente gametas ou se dividem para produzir células que depois se tornam gametas. A meiose, portanto, tem um papel central na reprodução entre eucariontes. Sem ela, os



(a)



(b)

Fig. 2.8 ■ Citocinese em uma célula animal (a) e de uma planta (b). A célula animal é um zigoto, que está se dividindo pela primeira vez. A citocinese é feita por constrição da célula em divisão ao redor de sua metade. Essa constrição cria um sulco, que é visto aqui em um lado da célula em divisão. Nas células vegetais, a citocinese é feita pela formação de uma placa membranosa entre as células filhas. Posteriormente, são construídas paredes de celulose em ambos os lados da placa celular.

organismos duplicariam seu número de cromossomos a cada geração, uma situação que rapidamente seria insuportável em vista das óbvias limitações de tamanho e capacidade metabólica das células.

Se observarmos os cromossomos em uma célula diplóide, veremos que eles estão aos pares (Fig. 2.9). Por exemplo, as células somáticas humanas têm 23 pares de cromossomos. Cada par é distinto. Pares diferentes de cromossomos levam conjuntos diferentes de genes. Os membros de um par são chamados cromossomos homólogos, ou simplesmente **homólogos**, de uma palavra grega que significa “em concordância com”. Os homólogos levam os mesmos conjuntos de genes, embora, como veremos no Cap. 5, eles possam levar alelos diferentes desses genes. Cromossomos de pares dife-

rentes são chamados **heterólogos**. Durante a meiose, os homólogos se associam intimamente uns com os outros. Essa associação é a base de um processo ordenado que finalmente reduz o número de cromossomos a um estado haplóide. A redução do número de cromossomos ocorre de tal modo que cada uma das células haplóides resultantes recebe exatamente um membro de cada par cromossômico.

O processo de meiose envolve duas divisões celulares (Fig. 2.10). A duplicação cromossômica, que está associada à síntese de DNA, ocorre antes da primeira destas divisões. Ela não ocorre entre as duas divisões. Assim, a progressão de eventos é: duplicação cromossômica → divisão I da meiose → divisão II da meiose.

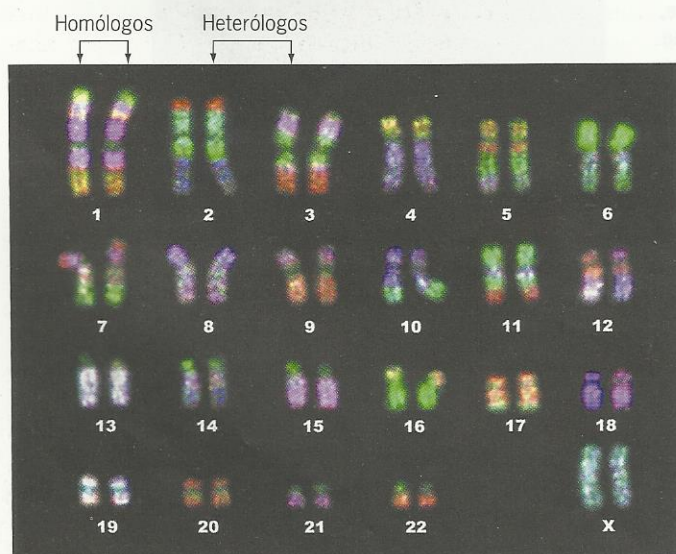


Fig. 2.9 ■ Os 23 pares de cromossomos homólogos encontrados em células humanas.

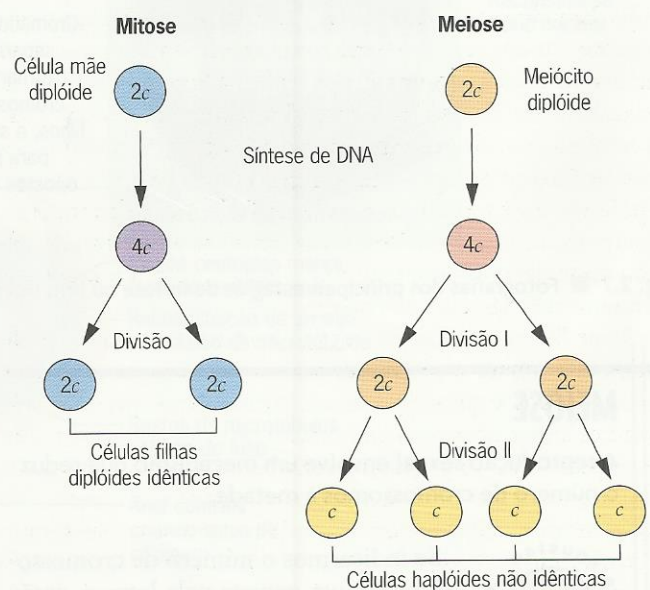


Fig. 2.10 ■ Comparação entre mitose e meiose; c é a quantidade haplóide de DNA no genoma.

Meiose I

Os eventos das duas divisões meióticas são ilustrados na Fig. 2.11 e na Fig. 2.12. A primeira divisão meiótica é complicada e prolongada. Quando ela começa, os cromossomos já estão duplicados; conseqüentemente, cada um deles consiste em duas cromátides irmãs. A prófase I da meiose – ou simplesmente **prófase I** – é dividida em cinco estágios, cada um indicado por um termo grego. Esses termos indicam características importantes sobre o aspecto ou o comportamento dos cromossomos.

Leptóteno, das palavras gregas que significam “filamentos finos”, é o primeiro estágio da prófase I. Durante o leptóteno (também chamado estágio de leptonema), os cromossomos duplicados condensam-se no emaranhado difuso da cromatina. Ao microscópio óptico, cromossomos individuais são pouco distinguíveis. Com um microscópio eletrônico, cada um dos cromossomos parece consistir em duas cromátides irmãs. À medida que a condensação cromossômica continua, a célula progride para o **zigóteno** (de palavras gregas que indicam “filamentos pareados”). Durante o zigóteno (também chamado zigonema), cromossomos homólogos juntam-se intimamente. Este processo de pareamento entre homólogos é chamado **sinapse**. Em algumas espécies, a sinapse começa pelas pontas dos cromossomos e se espalha para as regiões médias. A sinapse é acompanhada da formação de uma estrutura protéica entre os cromossomos pareados (Fig. 2.13). Esta estrutura, chamada **complexo sinaptonêmico**, consiste em três estruturas pareadas – uma associada a cada um dos cromossomos (denominadas estruturas laterais) e uma central entre elas (denominada elemento central) – e um grande número de estruturas do tipo degraus que juntam os elementos laterais ao elemento central. O papel do complexo sinaptonêmico no pareamento dos cromossomos e nos subsequentes eventos meióticos não está totalmente entendido. Em alguns tipos de células meióticas, ele nem mesmo aparece. Assim, ele pode não ser absolutamente essencial para o pareamento durante a prófase I. O processo pelo qual homólogos se juntam na prófase I também não é bem compreendido. Estudos recentes sugerem que os homólogos podem de fato começar a se parear cedo na meiose I, durante o leptóteno. Esse pareamento pode ser facilitado pela tendência dos cromossomos homólogos em permanecer na mesma região do núcleo durante a intérfase. Assim, os homólogos não precisam ir longe para se encontrar.

À medida que a sinapse progride, os cromossomos duplicados continuam a se condensar em volumes menores. Os cromossomos espessados que resultam deste processo são característicos do **paquíteno** (das palavras gregas “filamentos espessos”). No paquíteno (também chamado estágio do paquinema), cromossomos pareados podem facilmente ser vistos ao microscópio óptico. Cada par consiste em homólogos duplicados, os quais consistem em duas cromátides irmãs. Se contarmos os homólogos, o par é chamado **bivalente** de cromossomos, enquanto, se forem contados os filamentos, eles são chamados **tétrades** de cromátides. Du-

rante o paquíteno, os cromossomos pareados podem trocar material (Fig. 2.14). Exploraremos este fenômeno, chamado **crossing-over**, e suas conseqüências no Cap. 7. Aqui, é suficiente dizer que as cromátides irmãs individuais podem romper-se durante o paquíteno, e os pedaços quebrados podem ser trocados entre as cromátides de uma tétrade. A quebra e a reunião que ocorrem durante o crossing-over podem, portanto, resultar em recombinação de material genético entre os cromossomos pareados. O fato de que esses tipos de trocas ocorreram pode ser visto à medida que a célula progride para o estágio seguinte da meiose I, o **diplóteno** (das palavras gregas para “dois filamentos”). Durante o diplóteno (também chamado estágio do diplonema), os cromossomos pareados separam-se um pouco. Entretanto, eles permanecem em contato íntimo onde fizeram o **crossing**. Estes pontos de contato são chamados **quiasmas** (da palavra grega que significa “cruz”). Um exame atento dos quiasmas indica que cada um deles envolve apenas duas das quatro cromátides na tétrade. O estágio de diplóteno pode durar um longo tempo. Em mulheres, por exemplo, ele pode persistir por mais de 40 anos.

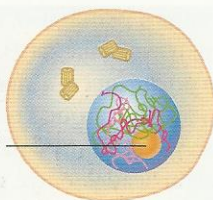
Próximo ao final da prófase I, os cromossomos condensam-se mais, a membrana nuclear fragmenta-se e um fuso acromático se forma. Os microtúbulos do fuso penetram no espaço nuclear e se ligam aos cinetócoros dos cromossomos. Os cromossomos, ainda mantidos juntos pelos quiasmas, movem-se para o plano central da célula que é perpendicular ao eixo do fuso. Esse movimento é característico do último estágio da prófase I, chamado **diacinese** (das palavras gregas indicando “movimento através”).

Durante a **metáfase I**, os cromossomos pareados orientam-se para pólos opostos do fuso. Essa orientação garante que, quando a célula se dividir, um membro de cada par irá para cada pólo. No final da prófase I e durante a metáfase I, os quiasmas que unem os bivalentes afastam-se dos centrômeros para as pontas dos cromossomos. Este fenômeno, chamado terminalização, reflete a crescente repulsão entre os membros de cada par cromossômico. Durante a **anáfase I**, os cromossomos pareados separam-se uns dos outros definitivamente. Essa separação, chamada **disjunção cromossômica**, é mediada pelo fuso que age em cada um dos bivalentes da célula. À medida que os cromossomos que se separam atingem os pólos opostos, a primeira divisão meiótica chega a um fim. Durante o estágio seguinte, chamado **telófase I**, o fuso é desfeito, as células filhas são separadas umas das outras por membranas, os cromossomos se descondensam, e um núcleo é formado ao redor dos cromossomos em cada uma das células filhas. Em algumas espécies, a descondensação dos cromossomos é incompleta, os núcleos filhos não se formam e as células filhas entram imediatamente na segunda divisão meiótica. As células produzidas pela meiose I contêm o número haplóide de cromossomos. Entretanto, cada cromossomo ainda consiste em duas cromátides irmãs, que podem não ser geneticamente idênticas porque podem ter trocado material com seus parceiros durante a prófase I.

Prófase I**Leptóteno**

Cromossomos replicados tornam-se visíveis.

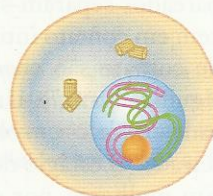
Núcléolo

**Zigóteno**

Pareamento de cromossomos homólogos.

**Paquíteno**

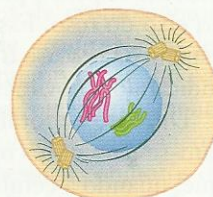
Cromossomos homólogos totalmente pareados. Ocorre crossing.

**Diplóteno**

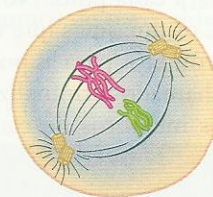
Cromossomos homólogos começam a se repelir. Cromátides tornam-se totalmente visíveis. Quiasmas tornam-se visíveis.

**Diacinese**

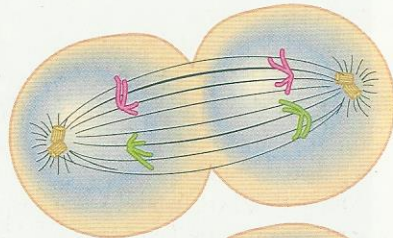
Cromossomos continuam a se encurtar e espessar. Núcléolo e membrana nuclear desaparecem. Microtúbulos ligam-se aos cinetócoros.

**Metáfase I**

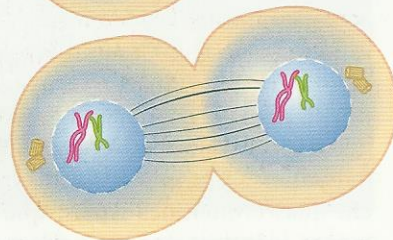
Montagem do fuso é completada. Cada par de cromossomos alinha-se na placa metafásica do fuso.

**Anáfase I**

Pares de cromossomos homólogos separam-se e migram para pólos opostos.

**Telófase I**

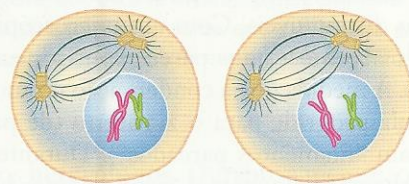
Cromossomos (cada com duas cromátides irmãs) completam migração para os pólos, e formam-se novas membranas nucleares.

**Citocinese**

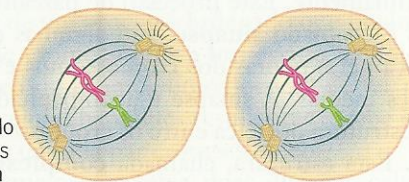
Na maioria das espécies, a citocinese produz duas células filhas. Cromossomos não se replicam antes da meiose II.

Prófase II

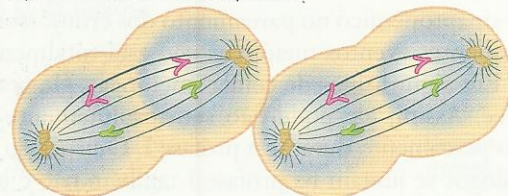
Cromossomos condensam-se e se movem para placa metafásica.

**Metáfase II**

Cinetócoros ligam-se às fibras do fuso. Cromossomos alinham-se na placa metafásica.

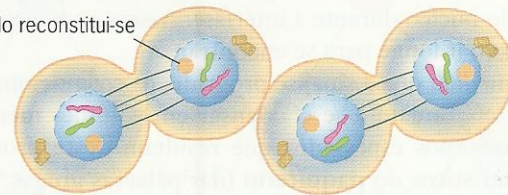
**Anáfase II**

Cromátides irmãs separam-se e se movem para pólos opostos à medida que cromossomos se separam.

**Telófase II**

Forma-se a membrana nuclear ao redor dos cromossomos e estes se deslicoidizam. Reconstitui-se o nucléolo.

Núcléolo reconstitui-se



Formam-se quatro células haplóides após citocinese.

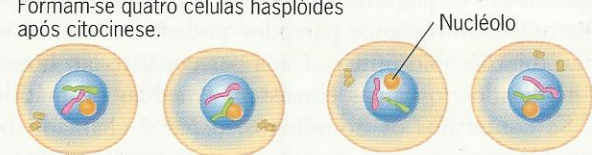


Fig. 2.11 ■ Meiose em animais.

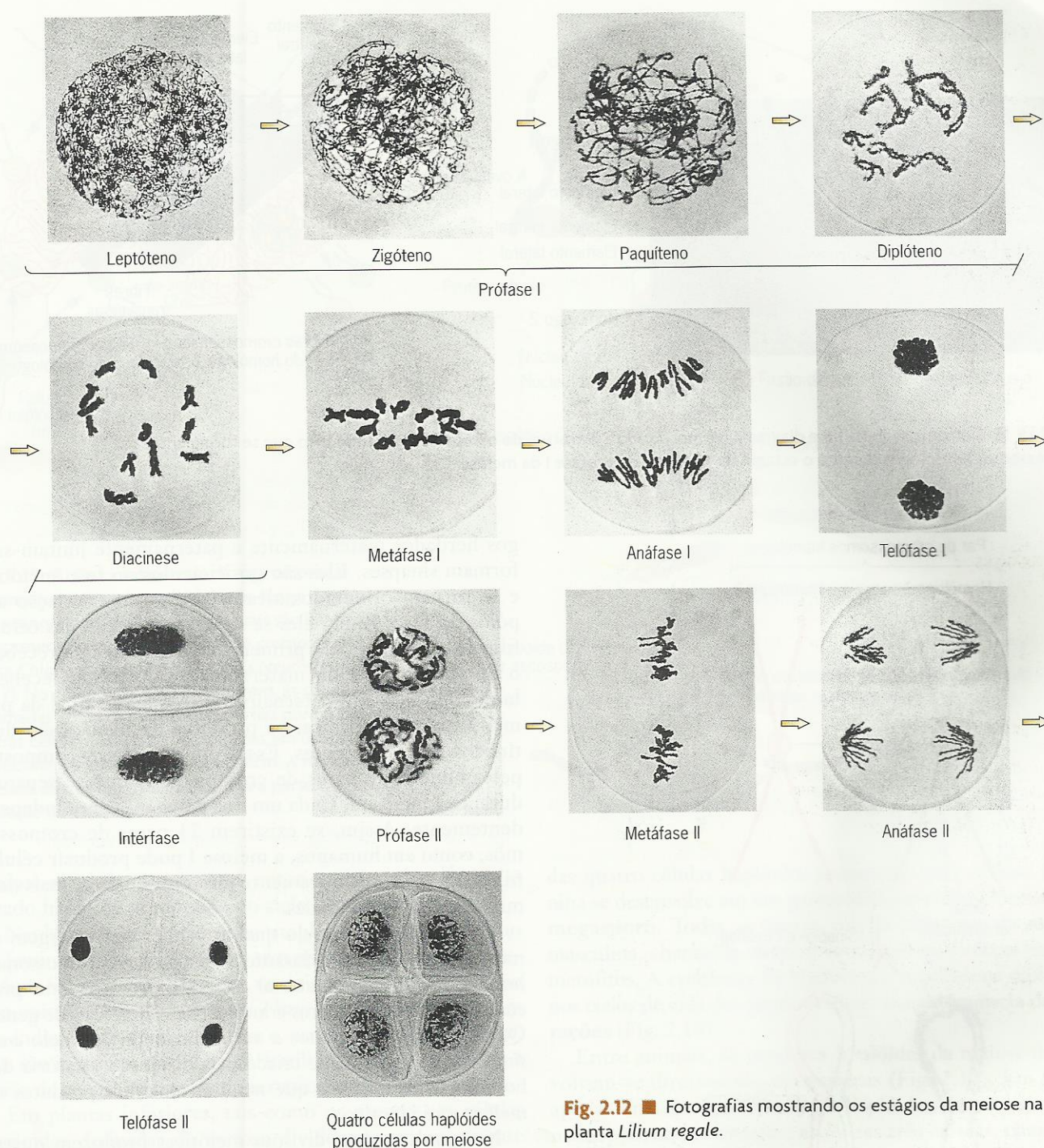


Fig. 2.12 ■ Fotografias mostrando os estágios da meiose na planta *Lilium regale*.

Meiose II e os Resultados da Meiose

Durante a meiose II, os cromossomos condensam-se e se ligam a um novo fuso acromático (**prófase II**). Eles então se movem para posições no plano equatorial da célula (**metáfase II**), e seus centrômeros se dividem para permitir que as cromátides irmãs se movam para pólos opostos (**anáfase II**), um fenômeno chamado **disjunção de cromátides**. Durante a **telófase II**, as cromátides separadas, agora chamadas cromossomos, atingem os pólos, e núcleos filhos se formam ao

redor deles. Cada núcleo filho contém um conjunto haplóide de cromossomos. Do ponto de vista do mecanismo, a meiose II parece muito uma mitose. Entretanto, seus produtos são haplóides, e, ao contrário dos produtos da mitose, as células que surgem da meiose II não são geneticamente idênticas.

Um motivo destas células diferirem é que os cromossomos homólogos formam pares e se separam uns dos outros durante a meiose I. Dentro de cada par de cromossomos, um homólogo foi herdado do organismo da mãe e o outro foi herdado de seu pai. Durante a meiose I, os homólo-

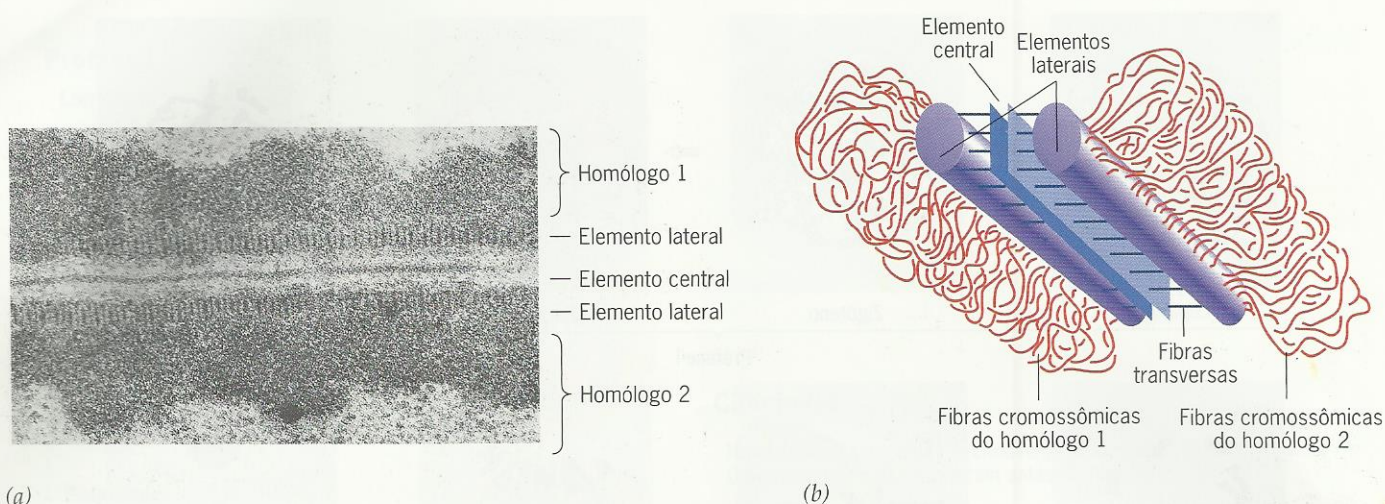


Fig. 2.13 ■ Eletromicrografia (a) e diagrama (b) mostrando a estrutura do complexo sinaptonêmico que se forma entre os cromossomos homólogos durante o estágio de zigóteno da prófase I da meiose.

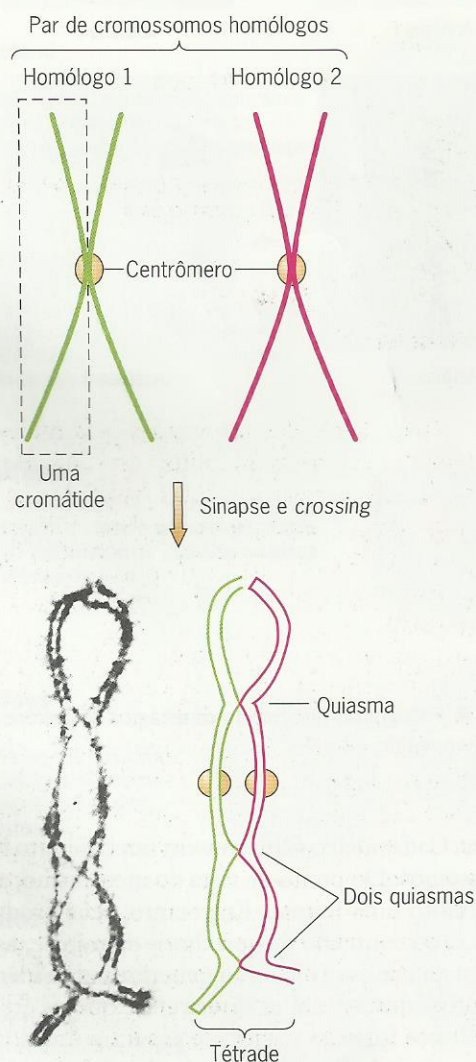


Fig. 2.14 ■ Quiasmas em um bivalente de cromossomos homólogos durante o estágio de diplóteno da prófase I da meiose.

gos herdados maternamente e paternamente juntam-se e formam sinapses. Eles são posicionados no fuso meiótico e se tornam orientados aleatoriamente com relação aos pólos do fuso. Então eles se separam. Metade das células filhas produzidas pela primeira divisão meiótica recebem o homólogo herdado maternamente e metade recebe o homólogo herdado paternamente. Assim, ao final da primeira divisão meiótica, os produtos da meiose são destinados a ser diferentes. Essas diferenças são compostas pelo número de pares de cromossomos que se separam durante a meiose I. Cada um dos pares se separa independentemente. Assim, se existirem 23 pares de cromossomos, como em humanos, a meiose I pode produzir células filhas 2^{23} cromossomicamente diferentes, isto é, mais de 8 milhões de possibilidades.

Um outro motivo pelo qual as células que emergem da meiose diferem é que, durante a meiose I, os cromossomos homólogos trocam material por *crossing-over*. Este processo pode criar inúmeras combinações diferentes de genes. Quando nós superpomos a variabilidade criada pelo *crossing-over* à variabilidade criada pela disjunção aleatória dos homólogos, é fácil ver que nenhum dos dois produtos da meiose será idêntico.

No geral, as duas divisões meióticas produzem quatro células, cada uma com o número haplóide de cromossomos. Tais células têm destinos diferentes em tipos diferentes de organismos. Em leveduras e algas unicelulares, elas se dividem mitoticamente para produzir populações de células haplóides. Em fungos simples, tais como o mofo do pão *Neurospora crassa*, elas se dividem mitoticamente para produzir filamentos de células chamados hifas, que formam o corpo do fungo. Em condições ambientais apropriadas, todos estes organismos haplóides podem produzir células sexuais, que são os seus gametas (Fig. 2.15). Quando duas células sexuais diferentes desses organismos se unem, elas produzem um zigoto diplóide, que geralmente entra em

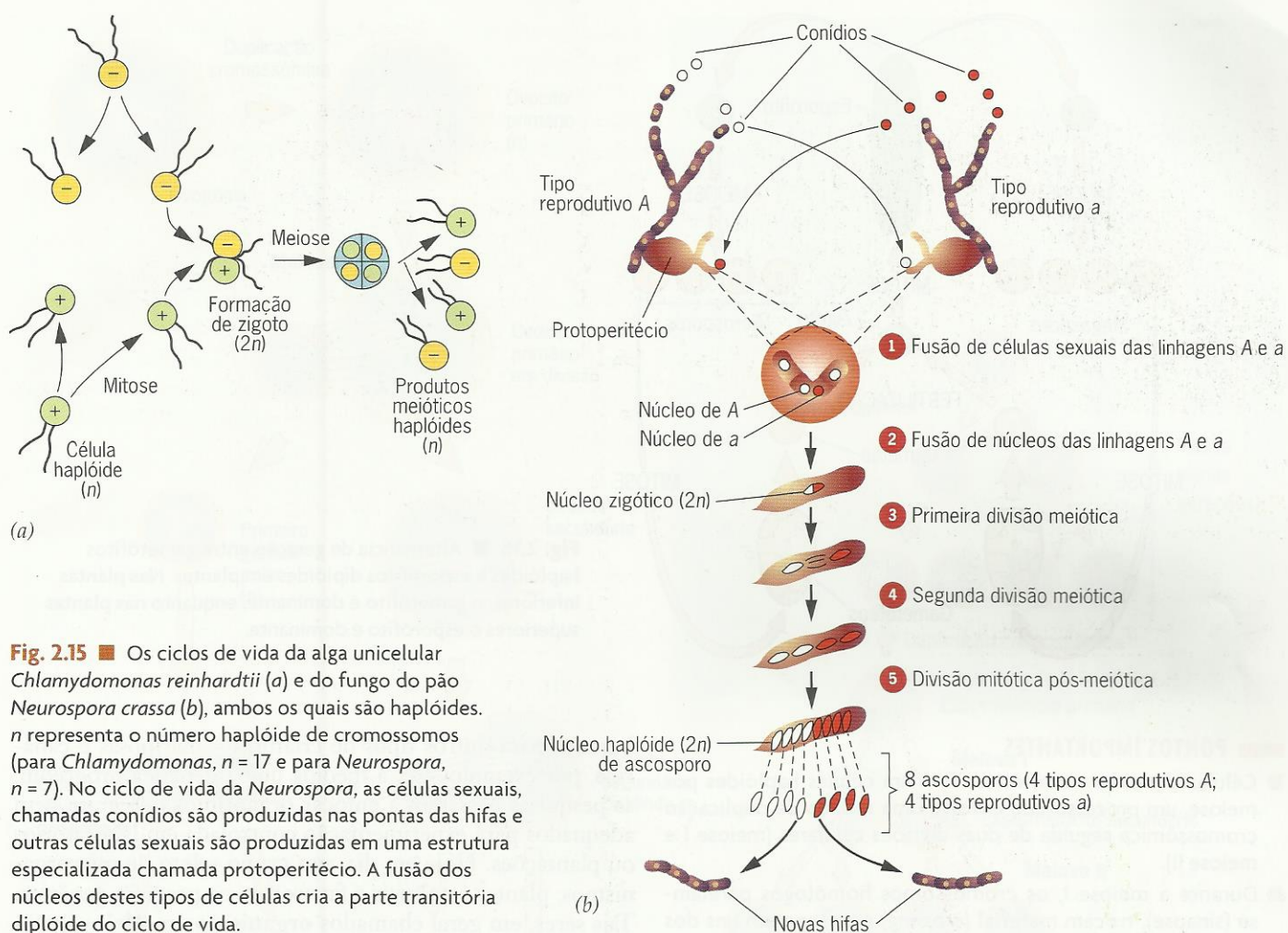


Fig. 2.15 ■ Os ciclos de vida da alga unicelular *Chlamydomonas reinhardtii* (a) e do fungo do pão *Neurospora crassa* (b), ambos os quais são haplóides. n representa o número haplóide de cromossomos (para *Chlamydomonas*, $n = 17$ e para *Neurospora*, $n = 7$). No ciclo de vida da *Neurospora*, as células sexuais, chamadas conídios são produzidas nas pontas das hifas e outras células sexuais são produzidas em uma estrutura especializada chamada protoperitécio. A fusão dos núcleos destes tipos de células cria a parte transitória diplóide do ciclo de vida.

meiose sem um crescimento mitótico intercalar. Assim, o estado haplóide predomina no ciclo de vida.

Nenhum destes organismos simples tem sexo masculino ou feminino diferenciado. Eles têm tipos reprodutivos. Na levedura *Saccharomyces cerevisiae*, os dois tipos reprodutivos são denominados *a* e *alfa*; na alga *Chlamydomonas reinhardtii*, eles são indicados por mais e menos; e, no bolor *Neurospora crassa*, eles são indicados por *A* e *a*. A fertilização sempre envolve a união de células sexuais de tipos reprodutivos opostos.

Em plantas inferiores, tais como os musgos, as células haplóides derivadas da meiose dividem-se mitoticamente para produzir filamentos ramificados que acabam por se diferenciar em tecidos tais como caules e folhas. Quando maduras, essas plantas haplóides produzem gametas, sejam ovócitos ou espermatozóides. Conseqüentemente, elas são chamadas **gametófitos**. Os gametas unem-se na fertilização para formar zigotos diplóides, que se dividem mitoticamente e se desenvolvem em estruturas chamadas **esporófitos**. Em plantas superiores, o gametófito é muito reduzido de tamanho e consiste em apenas algumas células haplóides. Assim, o esporófito é a parte conspícua do ciclo de vida. A meiose ocorre em tecidos reprodutivos masculinos e femininos distintos nos esporófitos de plantas superiores. Apenas uma

das quatro células haplóides produzidas pela meiose feminina se desenvolve em um gametófito; esta célula é chamada **megaspório**. Todas as quatro células haplóides da meiose masculina, chamadas **microspórios**, se desenvolvem em gametófitos. A existência de organismos haplóides e diplóides nos ciclos de vida das plantas é chamada **alternância de gerações** (Fig. 2.16).

Entre animais, os produtos haplóides da meiose desenvolvem-se diretamente em gametas (Fig. 2.17). Em geral, apenas uma das quatro células haplóides da meiose feminina se torna um ovócito; as outras três células, chamadas glóbulos polares, degeneram. Em contraste, todas as quatro células haplóides masculinas produzidas pela meiose se desenvolvem em espermatozóides. A formação de gametas, um processo chamado **gametogênese**, ocorre nas gônadas de animais. A ovocitogênese, a formação de ovócitos, ocorre nos ovários, que são as gônadas femininas, e a espermatogênese, a formação de espermatozóides, ocorre nos testículos, que são as gônadas masculinas. Tais processos começam quando células diplóides indiferenciadas, chamadas ovogônias ou espermatogônias, sofrem meiose para produzir células haplóides. As células haplóides então diferenciam-se em gametas maduros.

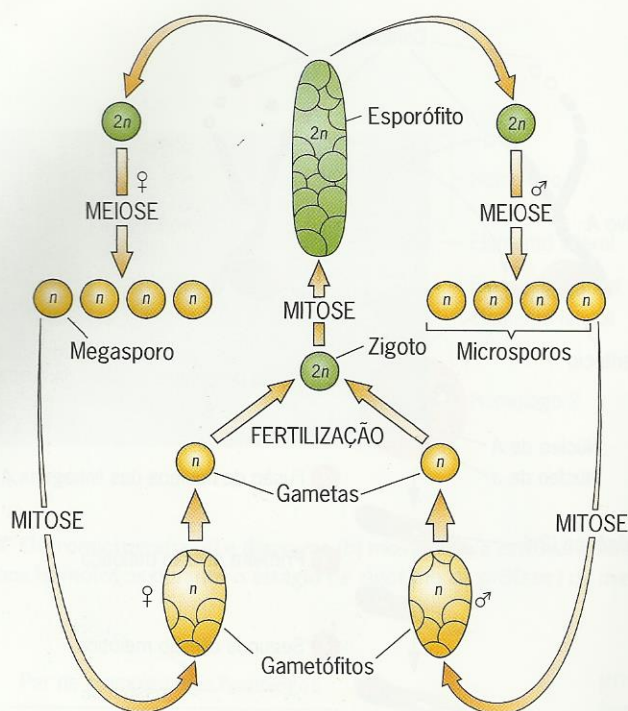


Fig. 2.16 ■ Alternância de geração entre gametófitos haplóides e esporófitos diplóides em plantas. Nas plantas inferiores, o gametófito é dominante, enquanto nas plantas superiores o esporófito é dominante.

PONTOS IMPORTANTES

- Células diplóides eucarióticas formam células haplóides por meiose, um processo que envolve uma rodada de duplicação cromossômica seguida de duas divisões celulares (meiose I e meiose II).
- Durante a meiose I, os cromossomos homólogos pareiam-se (sinapse), trocam material (*crossing*) e se separam uns dos outros.
- Durante a meiose II, as cromátides separam-se umas das outras.
- Em muitos organismos, os produtos haplóides da meiose desenvolvem-se diretamente em gametas.
- Em plantas, os produtos da meiose dividem-se mitoticamente para formar gametófitos haplóides.
- A fase gametofítica do ciclo de vida de uma planta alterna-se com a fase esporofítica, que é diplóide; a meiose ocorre no esporófito.

GENÉTICA NO LABORATÓRIO: UMA INTRODUÇÃO A ALGUNS ORGANISMOS MODELOS DE PESQUISAS

Os geneticistas enfocam sua pesquisa em microrganismos, plantas e animais bem adequados à experimentação.

Quando a genética começou, os organismos que eram usados para pesquisa eram os disponíveis em jardins ou terreiros. Alguns geneticistas pioneiros ampliaram o estudo da

herança para outros tipos de criaturas – mariposas e canários, por exemplo –, e, à medida que a genética progrediu, as pesquisas passaram a focar organismos que eram bem adequados para experimentação controlada em laboratórios ou plantações. Hoje em dia, um grupo seleto de microrganismos, plantas e animais é favorecido na pesquisa genética. Tais seres, em geral chamados **organismos modelos**, prestam-se bem à análise genética. Na maior parte, eles são facilmente cultivados no laboratório, seus ciclos de vida são relativamente curtos e eles são geneticamente variáveis. Além disso, pelos trabalhos ao longo de muitos anos, os geneticistas estabeleceram grandes coleções de linhagens mutantes para esses organismos. Encontraremos organismos genéticos modelos muitas vezes neste livro. Nas seções seguintes, destacaremos apenas alguns deles.

Escherichia coli, uma Bactéria

A análise genética de bactérias começou no início dos anos 1940. Inicialmente, várias espécies diferentes foram estudadas, mas em apenas alguns anos a maioria das pesquisas enfocou a *Escherichia coli*, uma bactéria em forma de bastonete que ocorre naturalmente nos intestinos de animais como nós mesmos. Este organismo pode ser cultivado no laboratório em um meio simples, presta-se a todos os tipos de análises bioquímicas e linhagens mutantes com diferentes requisitos de crescimento podem ser facilmente isoladas. Como resultado destas características, *E. coli* tornou-se o laboratório microbiano dos geneticistas. Muitas das primeiras descobertas em genética molecular vieram de estudos com *E. coli*. Por exemplo, o mecanismo

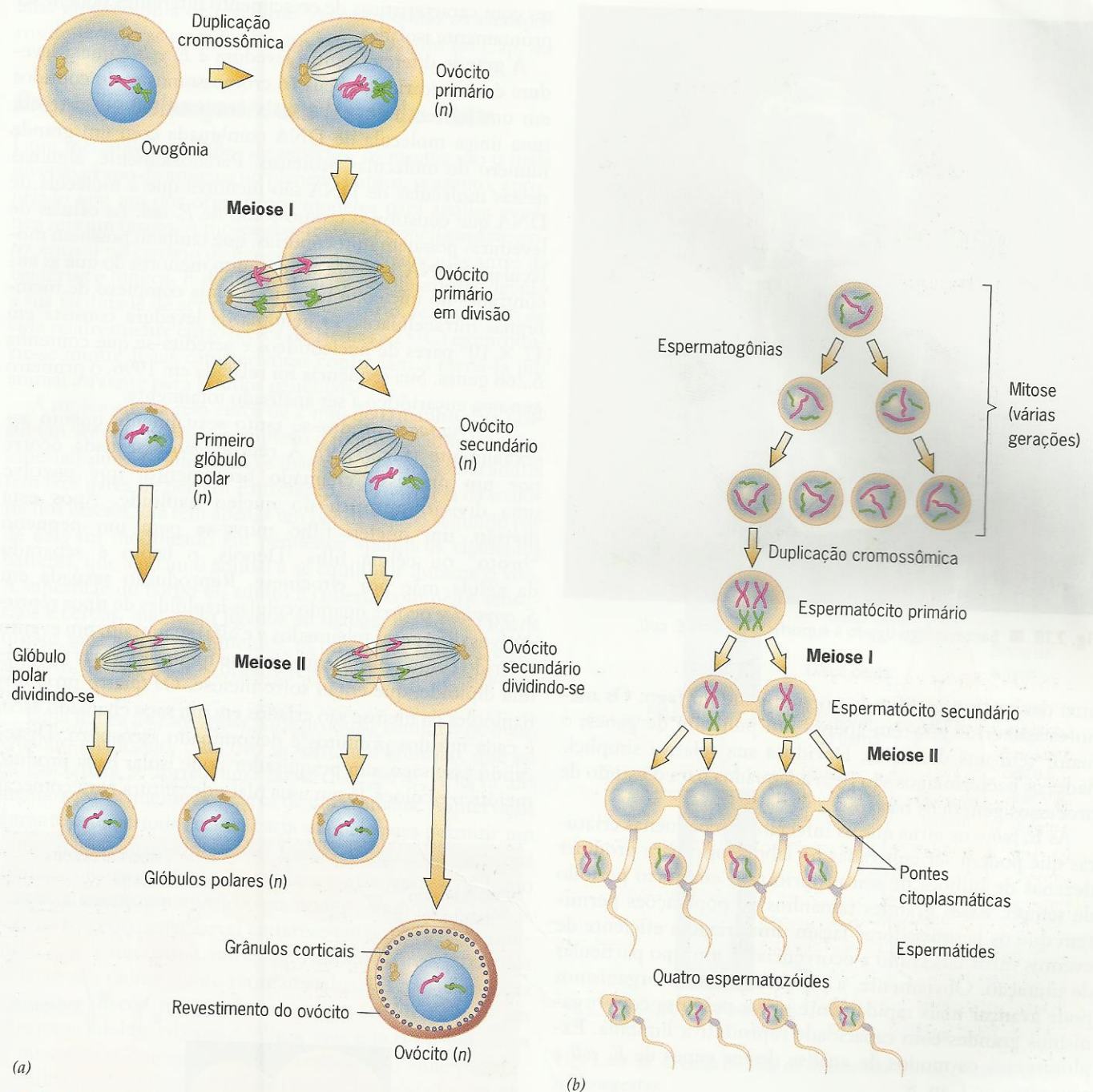


Fig. 2.17 ■ Gametogênese em animais. (a) A ovogênese em uma fêmea produz um ovócito e três glóbulos polares. (b) A espermatogênese em um macho produz quatro espermatozoides, que permanecem conectados uns aos outros por pontes citoplasmáticas até ficarem maduros.

de replicação do DNA e os processos de transcrição e tradução foram elucidados pelos trabalhos com esta bactéria. Veremos o papel que a *E. coli* desempenhou nestas descobertas nos Caps. 10, 11 e 12. O DNA do genoma de *E. coli* consiste em $4,6 \times 10^6$ pares de nucleotídeos dispostos em um único cromossomo circular com uma circunferência de 1,4 mm. A sequência completa deste genoma foi obtida em 1997. Sabe-se que ele contém 4.288 genes codificantes de proteína.

Muitas descobertas importantes em genética molecular também vieram de estudos com vírus que naturalmente infectam *E. coli*. Tais vírus, chamados **bacteriófagos**, da palavra grega que significa “comer bactéria”, são muito menores que as *E. coli*, cerca de um milésimo do volume. Eles se ligam às superfícies celulares e injetam seu DNA na célula, onde este DNA se replica e se expressa para produzir mais vírus (Fig. 2.18). Cada tipo de bacteriófago, ou fago, tem uma estrutura característica composta de proteínas, cada



Fig. 2.18 ■ Bacteriófago ligado à superfície de uma *E. coli*.

uma das quais é codificada por um gene do fago. Os menores bacteriófagos têm apenas um punhado de genes; o maior tem uns duzentos. Devido a sua relativa simplicidade, os bacteriófagos são bem adequados para o estudo de processos genéticos básicos.

As *E. coli* e os vírus que as infectam são pequenas criaturas que podem ser cultivadas no laboratório para produzir dezenas de bilhões de seu próprio tipo em curto período de tempo. Esses grandes tamanhos de populações permitem que os pesquisadores façam uma triagem eficiente de eventos raros tais como a ocorrência de um tipo particular de mutação. Obviamente, a pesquisa com tais organismos pode avançar mais rapidamente que a pesquisa com organismos grandes com capacidade reprodutiva limitada. Exploraremos os modos de análise destes genes de *E. coli* e seus vírus no Cap. 8.

Saccharomyces cerevisiae, a Levedura do Pão

A levedura do pão entrou na pesquisa genética mais ou menos na mesma época que a *E. coli*. Entretanto, muito antes de ser comum em laboratórios de genética, este organismo já era usado em culinária para fazer o pão. A levedura é um fungo unicelular, embora em algumas condições suas células dividam-se para formar longos filamentos, que se assemelham às hifas de outros fungos. Como a *E. coli*, as leveduras podem ser cultivadas em meios simples no laboratório, e são obtidos grandes números delas a partir de uma única célula mãe em apenas alguns dias. Além disso, linhagens mutan-

tes com características de crescimento diferentes podem ser prontamente isoladas.

A grande diferença entre levedura e *E. coli* é que a levedura é um eucarionte. Seus 16 cromossomos estão contidos em um núcleo definido, e cada cromossomo consiste em uma única molécula de DNA combinada com um grande número de moléculas protéicas. Paradoxalmente, algumas dessas moléculas de DNA são menores que a molécula de DNA que constitui o cromossomo de *E. coli*. As células de leveduras possuem mitocôndrias, que também possuem moléculas de DNA, embora sejam muito menores do que as encontradas no núcleo, e têm um sistema complexo de membranas intracelulares. O genoma da levedura consiste em 12×10^6 pares de nucleotídeos e acredita-se que contenha 6.268 genes. Sua sequência foi relatada em 1996, o primeiro genoma eucariótico a ser analisado totalmente.

S. cerevisiae reproduz-se tanto sexualmente quanto assexualmente (Fig. 2.19). A reprodução assexuada ocorre por um processo chamado brotamento, que envolve uma divisão mitótica do núcleo haplóide. Após esta divisão, um núcleo filho move-se para um pequeno "broto" ou célula filha. Depois, o broto é separado da célula mãe por citocinese. Reprodução sexual em *S. cerevisiae* ocorre quando células haplóides de tipos reprodutivos diferentes (chamados *a* e *α*) se juntam, um evento referido como cruzamento, para se fundir e formar uma célula diplóide, que então sofre meiose. Os quatro produtos haplóides da meiose são criados em um saco chamado asco, e cada um dos produtos é denominado ascosporo. Dissecando esse saco, um pesquisador pode isolar cada produto meiótico e colocá-lo em uma placa de cultura para começar

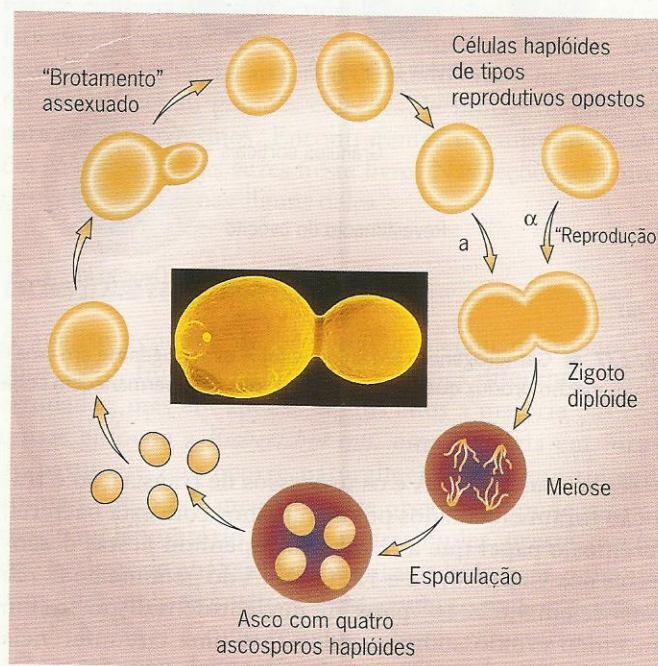


Fig. 2.19 ■ O ciclo de vida da levedura do pão, *Saccharomyces cerevisiae*. Um brotamento de levedura é mostrado no destaque.

uma nova colônia de levedura. Veremos como os ascospores de leveduras e outros fungos são analisados no Cap. 8.

Drosophila melanogaster, uma Mosca-das-frutas

Com incontáveis milhões de espécies, os insetos são o mais diverso grupo de animais no planeta. Parece, portanto, apropriado que um dos organismos modelos para análise genética seja um inseto. A mosca-das-frutas *Drosophila melanogaster* entrou no cenário da genética em 1909, décadas antes de bactérias e fungos. É de interesse para pesquisadores devido a sua facilidade de ser criada no laboratório, ter um ciclo de vida relativamente curto (cerca de 10 dias a 25°C) e reproduzir-se muito. Estas características continuam a torná-la um animal atraente para estudos genéticos.

A mosca adulta tem cerca de 2 mm de tamanho. Ela possui um sistema nervoso complexo e muitos tecidos e órgãos especializados. Como todos os insetos, uma *Drosophila* adulta tem três pares de pernas. Entretanto, ao contrário da maioria dos insetos, ela tem apenas um par de asas; o segundo par de asas foi modificado em pequenos apêndices chamados halteres, que ajudam o animal a se equilibrar durante o voo. A superfície do corpo do adulto é coberta com pêlos sensoriais e cerdas, que estão ligados ao sistema nervoso. Outros órgãos sensoriais proeminentes, os olhos e as antenas, estão situados na cabeça. Os órgãos reprodutivos estão situados no abdome.

Uma fêmea de *Drosophila* pode produzir centenas de ovos. Quando fertilizado, cada ovo se torna um embrião, que sai da casca para se tornar uma **larva** (Fig. 2.20). A larva alimenta-se vorazmente por cerca de uma semana. A intervalos durante este período, ela descarta sua pele para permitir um aumento no tamanho do corpo. Cada estágio larvar entre tais mudas é chamado **instar**. O terceiro instar larval é o último estágio antes da metamorfose em **imago**, ou mosca adulta. A pele do terceiro instar larval endurece em um casulo, e o animal torna-se uma **pupa**, um estágio durante o qual os tecidos da larva são radicalmente reorganizados. Grupos de células chamados **discos imaginais** crescem e se diferenciam em estruturas adultas tais como olhos, asas e pernas. Em cerca de quatro dias, emerge uma mosca adulta do casulo pupal.

Milhares de linhagens mutantes de *Drosophila* foram obtidas nos quase cem anos em que este animal foi estudado. Algumas apresentam cores de olhos e corpo incomuns; outras têm uma anatomia anormal – asas curtas, pernas extras, cerdas deslocadas, e assim em diante. Em uma linhagem, os halteres são convertidos em asas. A complexidade anatômica da *Drosophila* proporciona aos geneticistas a oportunidade de estudar a maneira como os genes controlam a formação das partes do corpo. Também lhes permite investigar os processos pelos quais um zigoto se desenvolve em um embrião e depois em um adulto. Um outro invertebrado, o nematódeo *Caenorhabditis elegans*, conhecido pelos geneticistas simplesmente como *C. elegans*, também desempenhou um importante papel na análise genética do desenvolvimento.



Fig. 2.20 ■ O ciclo de vida da mosca-das-frutas, *Drosophila melanogaster*.

Exploraremos algumas das descobertas deste ramo da genética no Cap. 22.

Tanto o genoma de *Drosophila* quanto o de *C. elegans* foram seqüenciados. Estes genomas são muito maiores que o genoma de levedura – 100×10^6 pares de nucleotídeos para *C. elegans* e 170×10^6 pares de nucleotídeos para *Drosophila* – e contêm mais genes, uma estimativa de 20.516 em *C. elegans* e 13.792 em *Drosophila*. Entretanto, o número de genes nesses animais multicelulares é apenas duas a três vezes maior que o número de genes de levedura. Os geneticistas estão tentando associar este relativamente modesto aumento no número de genes ao incrível aumento de complexidade biológica.

Arabidopsis thaliana, Uma Planta de Crescimento Rápido

Captando a energia solar e a estocando em açúcares simples tais como glicose, as plantas desempenham um papel indispensável na rede da vida. Como subproduto deste processo, elas também geram oxigênio, que nós e todos os outros organismos não fotossintéticos respiramos. Plantas de jardim foram os primeiros organismos a ser estudados geneticamente. Hoje em dia, os geneticistas enfocam sua atenção na *Arabidopsis thaliana*, uma erva às vezes chamada agrião caseiro. Este membro da família das crucíferas é relacionado a plantas alimentícias tais como rabanete, couve e canola. Entretanto, não tem valor agrônômico ou horticultural.

Arabidopsis é pequena e tem um período de geração relativamente curto (cerca de cinco semanas). Como as ervilhas de jardim de Mendel, *Arabidopsis* é uma espécie autotélica; entretanto, linhagens diferentes podem ter fertilização cruzada no laboratório para produzir híbridos. Os gametas masculinos, ou microsporos, são produzidos pelas células mães dos microsporos por meiose nas anteras, que estão em

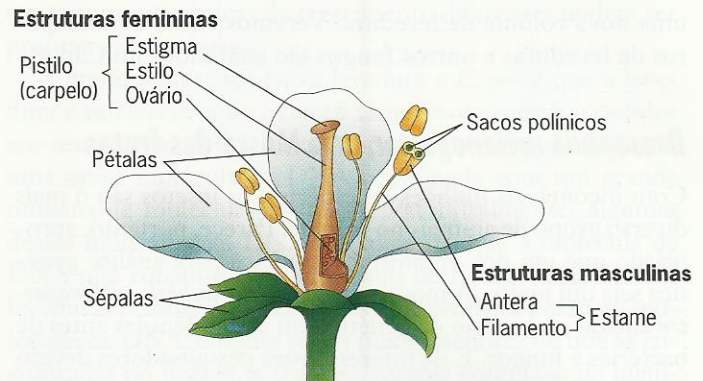


Fig. 2.21 ■ Órgãos reprodutivos masculino e feminino em uma flor típica.

cima nos **estames** dentro das flores de *Arabidopsis* (Figs. 2.21 e 2.22). Cada microsporo então sofre mitose para produzir um grão de pólen maduro, que contém duas células geradoras ou espermatozoides situados dentro de uma célula vegetativa. Cada uma destas células contém um núcleo haplóide. Os gametas femininos, ou megasporos, são produzidos por células

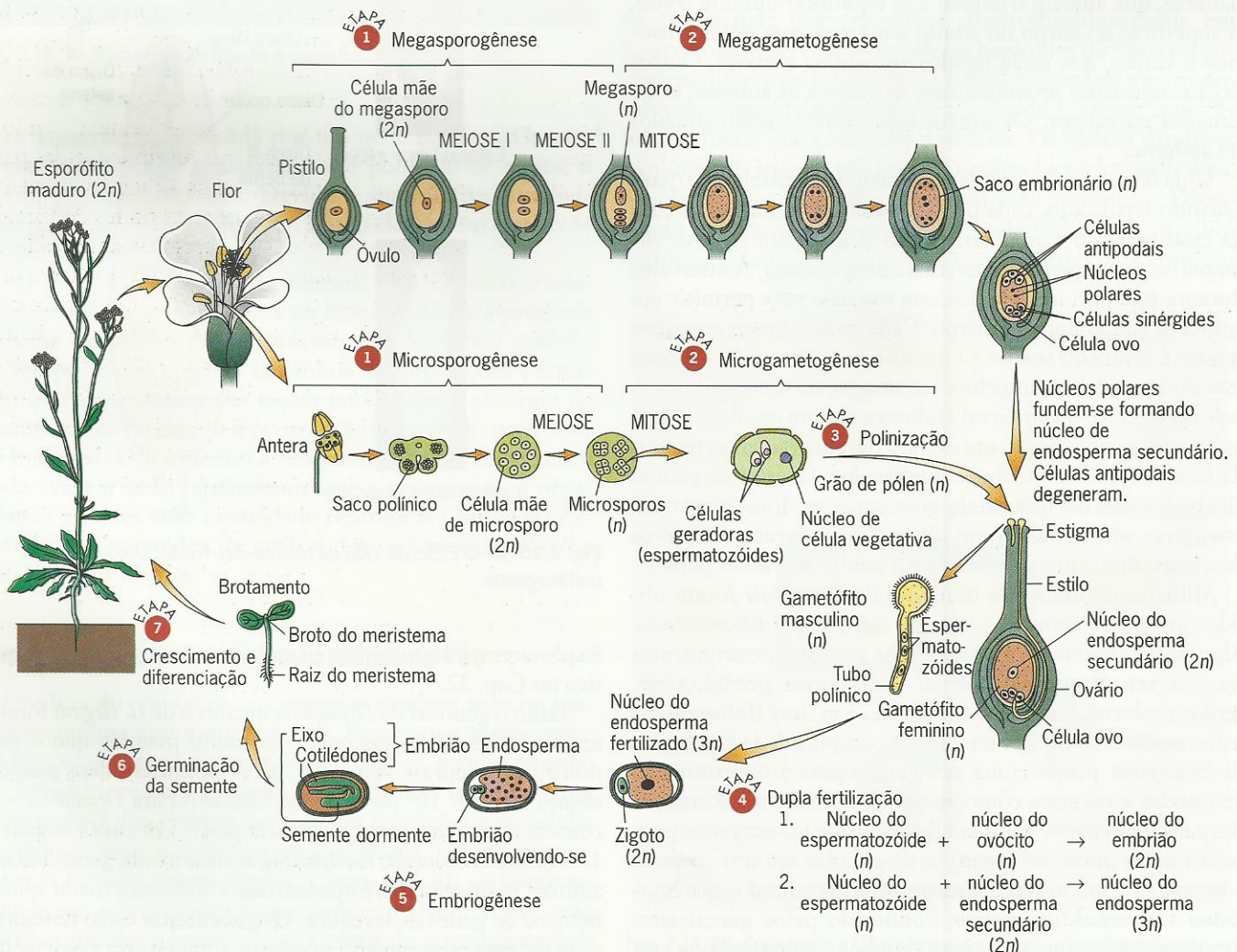


Fig. 2.22 ■ O ciclo de vida da planta modelo, *Arabidopsis thaliana*.

mães dos megasporos por meiose no **ovário**, que está situado dentro do **pistilo** no centro da flor. A meiose neste tecido feminino produz quatro células; entretanto, três subsequentemente degeneram, deixando apenas um megasporo funcional. O núcleo haplóide em cada megasporo sofre três divisões mitóticas para produzir um saco embrionário imaturo contendo oito núcleos haplóides. Ocorre então citocinese, criando três células antipodais, duas células sinérgide e uma célula ovo. Dois núcleos polares permanecem na grande célula central do saco embrionário. Esses núcleos polares subsequentemente se fundem para formar um **núcleo de endosperma secundário** diplóide, e as três células antipodais degeneram.

Quando um grão de pólen maduro cai no estigma em cima do pistilo, cresce um tubo polínico através do estilo até a célula ovo dentro do ovário. Em plantas, a fertilização envolve dois eventos. (1) O zigoto diplóide, que irá crescer em um embrião, é formado quando um espermatozóide do tubo polínico se funde com um ovócito do gametófito feminino. (2) O núcleo do endosperma triplóide é formado quando o outro núcleo do espermatozóide se combina com o núcleo diplóide do endosperma secundário do gametófito feminino. O tecido do endosperma triplóide, que se desenvolve a partir deste evento secundário de fertilização, alimenta o embrião enquanto a semente está germinando.

O genoma de *Arabidopsis* consiste em 157×10^6 pares de nucleotídeos de DNA, e acredita-se que contenha 27.706 genes. Muitos destes genes estão presentes em outras espécies de plantas, incluindo cultivos significativos do ponto de vista da agricultura tais como milho, arroz e trigo. Assim, a análise genética de *Arabidopsis* está fornecendo um arcabouço para compreender como os genes funcionam em plantas alimentícias importantes. Esta análise já produziu várias informações sobre o desenvolvimento de plantas, resistência a doenças, fotossíntese, tolerância ao frio e muitos outros fenômenos.

Homo sapiens, nossa Própria Espécie

O estudo da hereditariedade em seres humanos começou no século dezenove com análises estatísticas de características tais como altura e peso em genitores e proles. Em 1900, quando as descobertas de Mendel vieram à luz, a ênfase mudou para estudar a herança de características distintas e simples em famílias. Em geral, tais características eram doenças como a fenilcetonúria, que surge do metabolismo defeituoso de aminoácidos, ou a hemofilia, que surge de uma falha no mecanismo de coagulação do sangue. Com este enfoque, muitos genes humanos foram identificados, mas o progresso foi compreensivelmente lento. Seres humanos não podem ser submetidos a experimentação genética como as *Drosophila* ou leveduras. Neste aspecto, o *Homo sapiens* não é um organismo modelo. Entretanto, o desejo de aprender mais sobre o material genético de nossa própria espécie é muito forte, e grandes esforços têm sido feitos para satisfazer este desejo a despeito de muitos obstáculos.

Uma conquista surgiu em meados do século vinte, quando os biólogos aprenderam a multiplicar células humanas em cultura. Veja Um Marco na Genética: Cultura de Células Humanas. Tais células podem ser manipuladas em todo tipo de experimentos. Outra conquista surgiu no último quarto do século vinte, quando os pesquisadores aprenderam como isolar segmentos do DNA humano e propagá-los dentro de *E. coli*, uma técnica chamada **clonagem do DNA**. Finalmente, essa tecnologia abriu caminho para o Projeto do Genoma Humano. A análise do DNA humano também foi ajudada por estudos paralelos do DNA de outras espécies de vertebrados, incluindo o camundongo *Mus musculus* e o peixe-zebra *Danio rerio*. Estes dois organismos modelos podem ser submetidos a experimentação genética, e muitas das descobertas que vieram do trabalho com eles são relevantes para investigações genéticas de nossa própria espécie.

O camundongo é especialmente importante por causa de seu amplo emprego em muitas áreas da pesquisa biomédica. Camundongos têm sido usados para estudar a saúde e a doença e foram objeto de inúmeros projetos para avaliar os efeitos de drogas, substâncias químicas, alimentos e outros materiais relevantes para a saúde humana. A genética de camundongos começou cedo no século vinte com estudos sobre a herança da cor da pelagem. Desde aqueles dias, desenvolveu-se em um empreendimento notável. Foram produzidas linhagens especiais por cruzamento intenso para fornecer, aos pesquisadores, animais que são geneticamente uniformes, e milhares de linhagens mutantes foram coletadas para identificar genes e mapear cromossomos. O sequenciamento do genoma de camundongos forneceu um ponto de partida do qual os pesquisadores podem ter acesso ao significado de seqüências no genoma humano. As linhagens que deram origem a camundongos e humanos divergiram de um ancestral comum há cerca de 80 milhões de anos. Se uma seqüência tiver sido conservada nestes dois genomas durante todo este tempo evolutivo, ela deve ter uma função importante. Assim, a comparação dos genomas de camundongos e humanos, a prática da genômica comparativa, fornece um meio para avaliar que seqüências são funcionalmente importantes.

O genoma humano não é o maior ou o mais complexo genoma do planeta. Seus $3,2 \times 10^9$ pares de nucleotídeos contêm de 20.000 a 25.000 genes, bem como muito DNA não gênico. O Projeto do Genoma Humano, aliado a outros projetos de sequenciamento genômico, tem fornecido um arcabouço para analisar este material, mas muitos esforços ainda são necessários para esclarecer detalhes.

PONTOS IMPORTANTES

- A bactéria *E. coli* foi o primeiro procarionte da análise genética.
- Os eucariontes modelos incluem leveduras (*S. cerevisiae*), uma mosca-das-frutas (*D. melanogaster*) e uma planta de crescimento rápido (*A. thaliana*).
- Técnicas tais como cultura de células e clonagem do DNA possibilitaram o estudo do material genético de seres humanos e muitos outros organismos.



UM MARCO NA GENÉTICA: Cultura de Células Humanas

Os organismos modelos de genética são pequenos e têm ciclos de vida curtos. Como resultado, eles podem ser criados com baixo custo e em grandes números, e várias gerações – genitores, prole, netos etc. – podem ser examinadas em semanas ou meses, e não em anos. A facilidade e capacidade de estudar membros de gerações diferentes em um período de tempo razoável facilitam muito o progresso em pesquisas genéticas.

Ao contrário de organismos modelos, nossa própria espécie é grande e tem um longo ciclo de vida. Qualquer pessoa que queira fazer uma pesquisa genética sobre seres humanos terá que lidar com dificuldades e custos associados a aspectos humanos, bem como com o longo tempo necessário para obter resultados. As perspectivas de sucesso neste tipo de pesquisa não são muito boas, e além disso poucas pessoas estão dispostas a isto, seja como pesquisadores seja como objeto de pesquisa. Pesquisas envolvendo aspectos humanos são reguladas por princípios morais e éticos. Às vezes, os cientistas ultrapassam os limites estabelecidos por estes princípios, e, quando o fazem, nós os julgamos aberrantes, dementes ou malvados. Josef Mengele, o médico que fez pesquisas em internos dos campos de concentração nazista, é um exemplo infame. Felizmente, os cientistas foram capazes de criar meios para pesquisar fenômenos genéticos fora do organismo humano. Tais enfoques *in vitro* usam células humanas em cultura.

Técnicas de cultura de células foram desenvolvidas durante a primeira metade do século vinte. Durante várias décadas, os pesquisadores aprenderam a cultivar células de uma variedade de organismos, incluindo camundongos, galinhas e outros vertebrados, em tubos de cultura. Foram necessários meios complexos que continham soro sanguíneo ou líquidos extraídos de embriões, juntamente com um suprimento de sais, para cultivar tais células *in vitro*. Foram feitas muitas tentativas de cultivar células humanas, mas o sucesso foi limitado. Os melhores resultados foram obtidos quando as células foram derivadas de cânceres humanos. Por exemplo, em 1951, George Gey, um professor da Johns Hopkins School of Medicine, estabeleceu uma cultura de células cervicais cancerosas de uma mulher com 30 anos de idade chamada Henrietta Lacks. Essas células multiplicaram-se no meio de cultura que Gey usou. Foram colhidas amostras da cultura para estabelecer subculturas, e em apenas alguns anos a linhagem de células HeLa (Fig. 1) crescia em laboratórios de todo o mundo. Como disse Gey em uma palestra que fez em 1955: "... a linhagem do carcinoma HeLa cresce abundantemente e foi mantida por mais de quatro anos... Foi cultivada em nosso laboratório e em muitos outros pelo mundo em uma variedade de soro diluído com e sem suplementos sintéticos. Seu amplo uso criou novas oportunidades no estudo de células cancerosas humanas e suas respostas a vários agentes".¹

¹Gey, G. O. 1954–1955. Some aspects of the constitution and behavior of normal and malignant cells maintained in continuous culture. *Harvey Lectures* 50: 154–229.

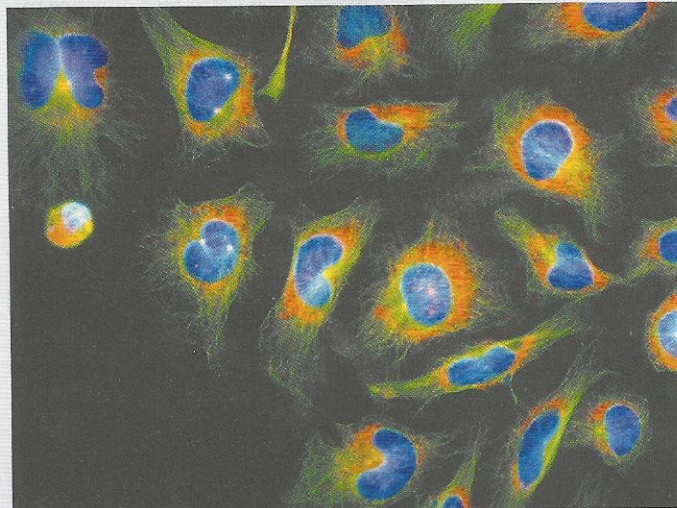


Fig. 1 ■ Células HeLa crescendo em cultura.

Muitos outros tipos de linhagens de células humanas foram então estabelecidos, a maioria, como a linhagem de células HeLa, derivadas de células cancerosas. Entretanto, também é possível cultivar células não cancerosas *in vitro*, embora tais células cresçam apenas por um tempo limitado. Tanto células humanas cancerosas quanto não cancerosas foram usadas em inúmeros experimentos para estudar estrutura, metabolismo, crescimento, divisão e comunicação celular, bem como o processo de diferenciação e desenvolvimento. Na genética, elas foram usadas para investigar a estrutura de cromossomos, localizar genes em cromossomos e analisar a expressão de genes e o funcionamento de seus produtos. Os geneticistas têm usado células humanas cultivadas para estudar fenômenos genéticos do mesmo modo que têm usado microrganismos tais como *E. coli* e leveduras.

Recentemente, pesquisadores começaram a enfocar sua atenção em modos de isolar e cultivar células tronco humanas. Essas células especiais podem ser obtidas de alguns tipos de tecidos adultos bem como de embriões. Com estimulação apropriada, elas podem ser induzidas a se diferenciar em cultura em tipos celulares especializados, tais como fibras musculares, linfócitos ou neurônios. Em virtude de sua capacidade de formar diversos tipos de células, as células tronco têm o potencial de produzir substituintes para tecidos desgastados e partes do corpo. Assim, elas estão nos noticiários. Entretanto, o modo pelo qual tais células são obtidas provocou controvérsias. Muitas pessoas questionam a destruição de embriões humanos para obter células tronco. Eles encaram este procedimento como imoral, e alguns o acham até similar ao trabalho do Dr. Mengele. Outras pessoas acreditam que é legítimo obter células tronco de embriões iniciais. Elas não questionam o procedimento e, mais, acreditam que verbas públicas devem estar disponíveis para financiar pesquisas com células tronco de origem embrionária. O debate continuado sobre células



UM MARCO NA GENÉTICA (continuação)

tronco indica que mesmo as pesquisas feitas fora do organismo humano podem criar importantes aspectos éticos. Para qualquer lado que os cientistas se voltem, eles encontrarão limites éticos.

QUESTÕES PARA DISCUSSÕES

1. Células não cancerosas podem ser cultivadas por 50 a 100 gerações. Elas então param de se dividir e acabam por morrer. As

células cancerosas podem ser mantidas indefinidamente em cultura. O que pode explicar as diferentes características de crescimento das células cancerosas e não cancerosas? O que o tempo finito de vida de culturas de células não cancerosas significa para o fenômeno de envelhecimento?

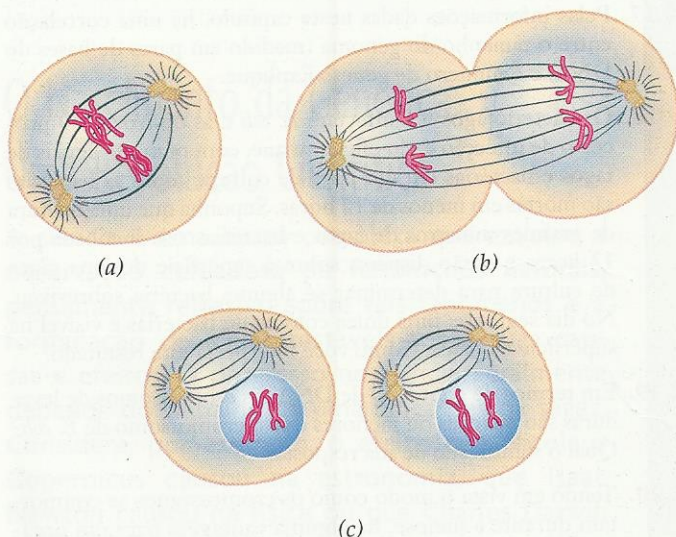
2. Quais são alguns dos potenciais usos terapêuticos das células tronco? Qual a evidência de que algumas dessas terapias potenciais podem de fato funcionar?

Exercícios Básicos Ilustram a análise genética básica.

1. Dentre os organismos modelos discutidos neste capítulo, quais são procariontes e quais são eucariontes? Os vírus enquadram-se neste esquema de classificação?

Resposta: *E. coli* é um procarionte; *S. cerevisiae*, *D. melanogaster*, *A. thaliana* e *H. sapiens* são todos eucariontes. Vírus são encontrados tanto em células procarióticas quanto eucarióticas; entretanto, em si, os vírus não são nem procariontes nem eucariontes.

2. Identifique os estágios de mitose nos seguintes desenhos.



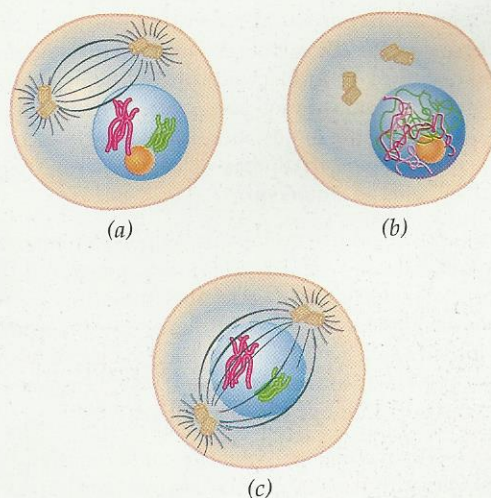
Resposta: (a) metáfase; (b) anáfase; (c) prófase.

3. Por que uma célula mãe diplóide que sofre meiose produz quatro células haplóides?

Resposta: Durante a meiose, a duplicação cromossômica precede dois eventos de divisão. Se o número de cromossomos

na célula mãe diplóide é $2n$, então após duplicação, a célula contém $4n$ cromátides. Durante a primeira divisão meiótica, os cromossomos homólogos formam pares e então são separados para células filhas diferentes, cada uma das quais recebe $2n$ cromátides. Durante a segunda divisão meiótica, o centrômero que une as duas cromátides de cada cromossomo se divide, e as cromátides são separadas para células filhas diferentes. Cada uma destas quatro células resultantes destas divisões meióticas sucessivas contém portanto n cromátides (agora chamadas cromossomos). Assim, o estado diplóide da célula mãe é reduzido ao estado haplóide nas quatro células que emergem da meiose.

4. Identifique os estágios da prófase I da meiose nos seguintes desenhos.



Resposta: (a) diplótenu; (b) leptótenu; (c) diacinese.

Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. Quais as principais diferenças entre mitose e meiose?

Resposta: Na mitose, um evento de divisão segue uma rodada de duplicação cromossômica. Na meiose, dois eventos de divisão seguem uma rodada de duplicação cromossômica. Além disso, durante a primeira divisão meiótica, os cromossomos homólogos pareiam-se. Esse pareamento baseado na homologia normalmente não ocorre durante a mitose. As duas células pro-

duzidas por uma divisão mitótica são idênticas uma à outra e à célula mãe da qual se derivaram. As quatro células produzidas pelas duas divisões meióticas sucessivas não são idênticas entre si nem à célula mãe da qual são derivadas. Quando uma célula diplóide sofre mitose, as duas células derivadas dela também serão diplóides. Quando uma célula diplóide sofre meiose, as quatro células derivadas dela serão haplóides.

Questões e Problemas Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 2.1. Carboidratos e proteínas são polímeros lineares. Que tipos de moléculas se combinam para formar estes polímeros?
- 2.2. Todas as células são circundadas por uma membrana; algumas células são circundadas por uma parede. Quais as diferenças entre membranas celulares e paredes celulares?
- 2.3. Quais as principais diferenças entre células procarióticas e eucarióticas?
- 2.4. Distinga os estados haplóide e diplóide. Que tipos de células são haplóides? Que tipos de células são diplóides?
- 2.5. Compare os tamanhos e estruturas dos cromossomos procarióticos e eucarióticos.
- 2.6. Enfocando os cromossomos, quais são os eventos principais durante a intérfase e a fase M no ciclo celular eucariótico?
- 2.7. Qual tipicamente dura mais, a intérfase ou a fase M? Você pode explicar por que uma destas fases dura mais que a outra?
- 2.8. De que modo diferem os centros organizadores de microtúbulos das células vegetais e animais?
- 2.9. Corresponda os estágios da mitose com os eventos que eles englobam: **Estágios:** (1) anáfase, (2) metáfase, (3) prófase, (4) telófase. **Eventos:** (a) reconstituição dos nucléolos, (b) desaparecimento da membrana nuclear, (c) condensação dos cromossomos, (d) formação do aparelho mitótico, (e) movimento dos cromossomos para a placa equatorial, (f) movimento dos cromossomos para os pólos, (g) descondensação dos cromossomos, (h) divisão do centrômero, (i) ligação dos microtúbulos ao cinetócoro.
- 2.10. Arrume os seguintes eventos na sequência temporal correta durante a divisão celular eucariótica, começando pela primeira: (a) condensação dos cromossomos, (b) movimento dos cromossomos para os pólos, (c) duplicação dos cromossomos, (d) formação da membrana nuclear, (e) ligação dos microtúbulos aos cinetócoros, (f) migração dos centrosomos para posições nos lados opostos do núcleo.
- 2.11. Um espermatozoide de *Drosophila* contém quatro cromossomos. Quantos cromossomos estariam presentes em uma espermatogônia ao entrar na meiose? Quantas cromátides estariam presentes em uma espermatogônia na metáfase I da meiose? Quantas estariam presentes na metáfase II?
- 2.12. Em seres humanos, o gene para β -globina está situado no cromossomo 11 e o gene para α -globina, que é outro componente da proteína hemoglobina, está situado no cromossomo 16. Seria esperado que estes dois cromossomos formassem par durante a meiose? Explique sua resposta.
- 2.13. O *crossing* ocorre antes ou depois da duplicação dos cromossomos nas células que estão sofrendo meiose?
- 2.14. Que características visíveis dos cromossomos indicam que eles sofreram *crossing* durante a meiose?
- 2.15. Durante a meiose, quando ocorre a disjunção de cromossomos? Quando ocorre a disjunção das cromátides?
- 2.16. Nas plantas com flores, o tecido esporofítico é haplóide ou diplóide? Quantos núcleos estão presentes no gametófito feminino? Quantos estão presentes no gametófito masculino? Eles são núcleos haplóides ou diplóides?
- 2.17. Pelas informações dadas neste capítulo, há uma correlação entre o tamanho do genoma (medido em pares de bases do DNA) e o número de genes? Explique.
- 2.18. Os bacteriófagos T1 infectam *E. coli* e as destroem. O processo de infecção é tão eficiente que, com números iguais de fagos e bactérias em um tubo de cultura, todas as bactérias são mortas em menos de 12 horas. Suponha que uma mistura de grandes números de fagos e bactérias seja incubada por 12 horas e então dispersa sobre a superfície de uma placa de cultura para determinar se alguma bactéria sobreviveu. No dia seguinte, uma única colônia de bactérias é visível na superfície da placa. Como você interpreta este resultado?
- 2.19. Em termos de conteúdo de DNA, os cromossomos de leveduras são maiores ou menores que o cromossomo de *E. coli*? Qual o significado de sua resposta?
- 2.20. Tendo em vista o modo como os cromossomos se comportam durante a meiose, há alguma vantagem para um organismo ter um número par de pares de cromossomos (como na *Drosophila*), em oposição a um número ímpar de pares de cromossomos (como nos seres humanos)?
- 2.21. Nas plantas com flores, dois núcleos do grão de pólen participam dos eventos de fertilização. Com quais núcleos do gametófito feminino estes núcleos se combinam? Que tecidos são formados pelos eventos de fertilização?

Mendelismo: Os Princípios Básicos da Herança

Tópicos

- Os Estudos de Mendel Sobre Hereditariedade
- Aplicações dos Princípios de Mendel
- Formulando e Testando Hipóteses Genéticas
- Princípios Mendelianos em Genética Humana

O Nascimento da Genética: Uma Revolução Científica

A ciência é uma busca complexa que envolve a observação cuidadosa de fenômenos naturais, pensamento reflexivo sobre tais fenômenos e a formulação de idéias testáveis sobre suas causas e efeitos. O progresso na ciência geralmente depende do trabalho de uma única pessoa sagaz. Considere, por exemplo, o efeito que Nicolaus Copernicus causou na astronomia, que Isaac Newton causou na física ou que Charles Darwin causou na biologia. Cada uma dessas pessoas alterou o curso de sua disciplina científica ao introduzir idéias radicalmente novas. Na verdade, eles deram origem a revoluções científicas.

Na metade do século dezenove, o monge austríaco Gregor Mendel, um contemporâneo de Darwin, fincou os alicerces de outra revolução na



Pisum sativum, o material dos experimentos de Gregor Mendel.

biologia, a que produziu uma ciência totalmente nova – a genética. As idéias de Mendel, publicadas em 1866 sob o título “Experimentos na Hibridização de Plantas”, eram destinadas a explicar como as características dos organismos são herdadas. Muitas pessoas tentaram uma explicação antes, mas sem muito sucesso. De fato, Mendel comentou sobre suas falhas nos parágrafos de abertura desse artigo:

Sobre este tema, vários observadores cuidadosos, tais como Kölreuter, Gärtner, Herbert, Lecoq, Wichura e outros, devotaram uma parte de suas vidas com perseverança incansável...

[Entretanto] Os que observarem o trabalho neste departamento chegarão à convicção de que, entre todos os numerosos experimentos feitos, nenhum teve esta amplitude e foi feito de tal modo, que possibilite a determinação do número de formas diferentes que a prole dos híbridos apresentou ou que arrumasse tais formas com certeza de acordo com suas gerações separadas ou que avaliasse definitivamente suas relações estatísticas.¹

Ele então descreveu seus próprios esforços para elucidar o mecanismo da hereditariedade.

De fato requer coragem para fazer um trabalho de tal alcance; isto parece, entretanto, ser o único modo certo pelo qual podemos finalmente alcançar a solução de uma dúvida cuja importância não pode ser subestimada quanto à conexão com a história da evolução das formas orgânicas.

Esta publicação agora apresentada registra os resultados de tal experimento detalhado. Este experimento foi praticamente confinado a um pequeno grupo de plantas e é agora, após oito anos de pesquisa, concluído em sua essência. Se o plano pelo qual foram conduzidos os experimentos separados era o mais adequado para atingir o fim desejado, é deixado para a amigável decisão do leitor.²

OS ESTUDOS DE MENDEL SOBRE HEREDITARIEDADE

Os experimentos de Gregor Mendel com ervilhas elucidaram o modo como as características são herdadas.



A vida de Gregor Johann Mendel (1822-1884) transpôs a metade do século dezenove. Seus genitores eram fazendeiros na Moravia, que então era parte do Império Hapsburg na Europa Central. Sua criação rural o

ensinou a plantar e a criar animais e inspirou seu interesse pela natureza. Aos 21 anos, Mendel deixou a fazenda e entrou para um monastério católico na cidade de Brünn (hoje, Brno na República Checa). Em 1847, ele foi ordenado padre, adotando o nome Gregor pela igreja. Depois, ele ensinou na escola local, tendo-se retirado de 1851 a 1853 para estudar na Universidade de Viena. Após retornar a Brünn, ele retomou sua vida como monge-professor

e começou seus experimentos genéticos que acabaram por torná-lo famoso.

Mendel fez experimentos com várias espécies de plantas de jardim e até tentou alguns experimentos com abelhas. Seu maior sucesso, entretanto, foi com ervilhas. Ele completou seus experimentos com ervilhas em 1864. Em 1865, Mendel apresentou os resultados à Natural History Society local e, no ano seguinte, publicou um relato detalhado nas publicações da sociedade. Veja Um Marco na Genética: Publicação de Mendel em 1866, no final deste capítulo. Infelizmente, esta publicação ficou na obscuridade até 1900, quando foi redescoberta por três botânicos, Hugo de Vries na Holanda, Carl Correns na Alemanha e Eric von Tschermak-Seysenegg na Áustria. À medida que estes homens pesquisaram a literatura científica à procura de dados que apoiassem suas próprias teorias de hereditariedade, cada um descobriu que Mendel já havia feito uma análise detalhada e cuidadosa há 35 anos. As idéias de Mendel rapidamente tiveram aceitação, sobretudo pelos esforços promocionais de um biólogo britânico, William Bateson. Este campeão das descobertas de Mendel criou um novo termo para descrever o estudo da hereditariedade: genética, da palavra grega que significa “gerar”.

^{1,2}Peters, J. A., ed. 1959. *Classic Papers in Genetics*. Prentice-Hall, Englewood Cliffs, NJ.

O Organismo Experimental de Mendel, a Ervilha de Jardim

Um motivo para o sucesso de Mendel é que ele escolheu astutamente seu material experimental. A ervilha de jardim, *Pisum sativum*, é uma dicotiledônea, um tipo de planta que produz duas folhas, ou cotilédones, a partir da germinação de uma semente. As ervilhas são facilmente cultivadas em jardins experimentais ou em vasos em uma estufa.

Uma peculiaridade da reprodução das ervilhas é que as pétalas da flor ficam bem fechadas, impedindo que os grãos de pólen entrem ou saiam. Isto força um sistema de autofertilização, no qual os gametas masculinos e femininos da mesma flor se unem para produzir sementes. Como resultado, as linhagens individuais de ervilhas são altamente endocruzadas, apresentando pouca ou nenhuma variação genética de uma geração para outra. Devido a esta uniformidade, dizemos que tais linhagens são **puras**.

Mendel obteve muitas variedades puras diferentes de ervilhas, cada uma distinta por uma determinada característica. Em uma linhagem, as plantas tinham 2 metros de altura, enquanto em outra elas mediam apenas meio metro. Outra variedade produzia sementes verdes, e ainda outra produzia sementes amarelas. Mendel tirou proveito destas características contrastantes para determinar como são herdadas tais características de ervilhas. Seu foco nestas diferenças únicas entre linhagens de ervilhas permitiu-lhe estudar a herança de uma característica de cada vez – por exemplo, a altura da planta. Outros biólogos tentaram seguir a herança de muitas características simultaneamente, mas, como os resultados de tais experimentos eram complexos, eles foram incapazes de descobrir qualquer princípio fundamental sobre hereditariedade. Mendel teve sucesso no contexto em que tais biólogos falharam porque enfocou sua atenção em diferenças contrastantes entre plantas que eram iguais no restante – alta *versus* baixa, sementes verdes *versus* amarelas, e assim em diante. Além disso, ele manteve registros cuidadosos dos experimentos que fez.

Cruzamentos Monoíbridos:

Os Princípios de Dominância e Segregação

Em um experimento, Mendel fez **fertilização cruzada** – ou, simplesmente, cruzou – plantas altas e baixas para investigar como a altura era herdada (Fig. 3.1). Ele removeu cuidadosamente as anteras de uma variedade antes que seu pólen tivesse amadurecido e então aplicou o pólen de outra variedade ao estigma, um órgão pegajoso no topo do pistilo que leva ao ovário. As sementes que resultaram destas fertilizações cruzadas foram semeadas no ano seguinte, produzindo híbridas que eram uniformemente altas. Mendel obteve plantas altas independente do modo como fazia o cruzamento (masculina alta com feminina baixa ou masculina baixa com feminina alta); assim, os dois cruzamentos recíprocos deram os mesmos resultados. Ainda mais significativo, Mendel notou que

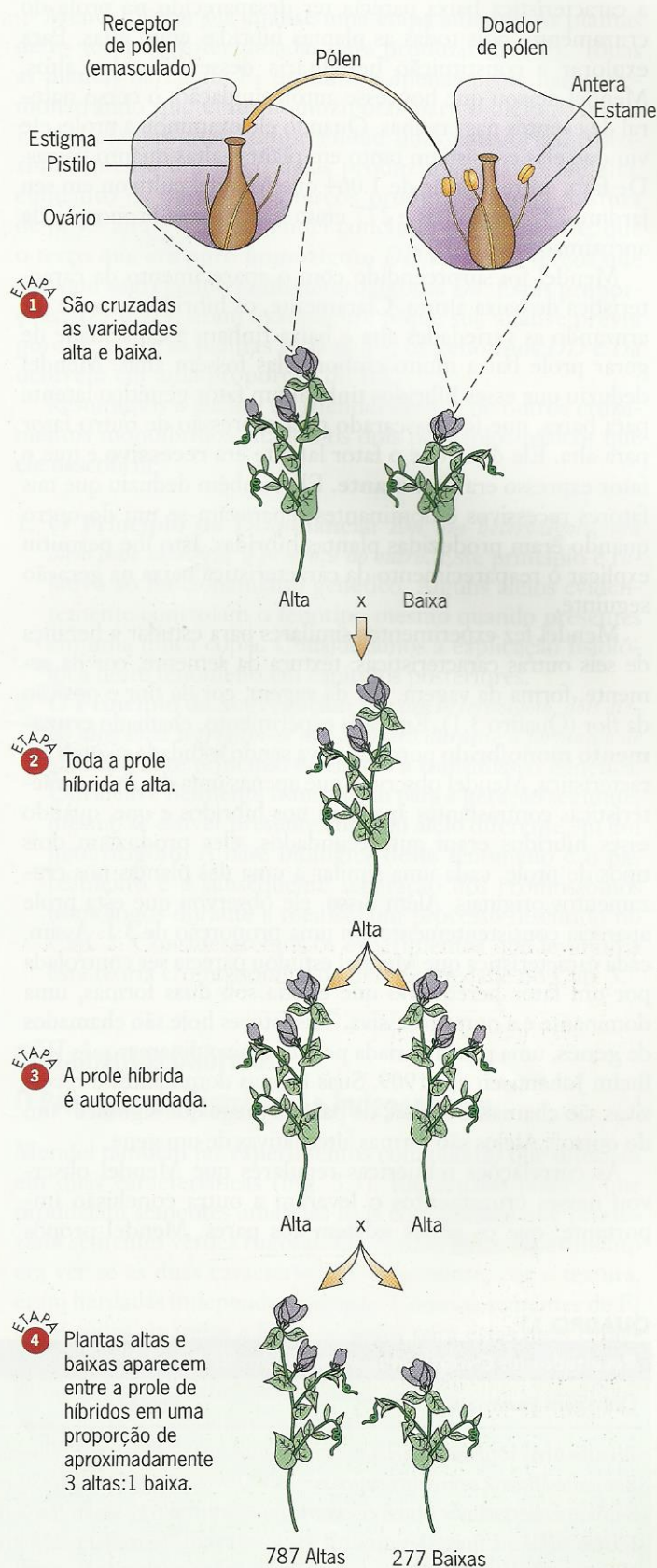


Fig. 3.1 ■ Os cruzamentos de Mendel envolvendo as variedades alta e baixa de ervilhas.

a característica baixa parecia ter desaparecido na prole do cruzamento, pois todas as plantas híbridas eram altas. Para explorar a constituição hereditária desses híbridos altos, Mendel deixou que houvesse autofecundação, o curso natural de eventos nas ervilhas. Quando ele examinou a prole, ele viu que elas consistiam tanto em plantas altas quanto baixas. De fato, entre a prole de 1.064 que Mendel cultivou em seu jardim, 787 eram altas e 277 eram baixas, uma proporção de aproximadamente 3:1.

Mendel foi surpreendido com o aparecimento da característica de baixa altura. Claramente, os híbridos que ele fez cruzando as variedades alta e baixa tinham a capacidade de gerar prole baixa muito embora elas fossem altas. Mendel deduziu que esses híbridos tinham um fator genético latente para baixa, que foi mascarado pela expressão de outro fator para alta. Ele disse que o fator latente era **recessivo** e que o fator expresso era **dominante**. Ele também deduziu que tais fatores recessivos e dominantes separavam-se um do outro quando eram produzidas plantas híbridas. Isto lhe permitiu explicar o reaparecimento da característica baixa na geração seguinte.

Mendel fez experimentos similares para estudar a herança de seis outras características: textura da semente, cor da semente, forma da vagem, cor da vagem, cor da flor e posição da flor (Quadro 3.1). Em cada experimento, chamado **cruzamento monoíbrido** porque estava sendo estudada só uma característica, Mendel observou que apenas uma das duas características contrastantes apareceu nos híbridos e que, quando esses híbridos eram autofecundados, eles produziam dois tipos de prole, cada uma similar a uma das plantas nos cruzamentos originais. Além disso, ele observou que esta prole aparecia consistentemente em uma proporção de 3:1. Assim, cada característica que Mendel estudou parecia ser controlada por um fator hereditário que existia sob duas formas, uma dominante e a outra recessiva. Tais fatores hoje são chamados de **genes**, uma palavra criada pelo botânico dinamarquês Wilhelm Johannsen em 1909. Suas formas dominantes e recessivas são chamadas **alelos**, da palavra grega que significa “um do outro”. Alelos são formas alternativas de um gene.

As correlações numéricas regulares que Mendel observou nesses cruzamentos o levaram a outra conclusão importante: que os genes existem aos pares. Mendel propôs

que cada uma das linhagens parentais que ele usou em seus experimentos levava duas cópias idênticas de um gene – na terminologia moderna, elas são diplóides e **homozigotas**. Entretanto, durante a produção de gametas, Mendel propôs que estas duas cópias são reduzidas a uma; isto é, os gametas que emergem da meiose levam uma única cópia de um gene – na terminologia moderna, eles são haplóides.

Mendel reconheceu que o número diplóide de genes seria restaurado quando o espermatozóide e o ovócito se uniam para formar um zigoto. Além disso, ele compreendeu que se o espermatozóide e o ovócito viessem de plantas geneticamente diferentes, como ele fez em seus cruzamentos, o zigoto híbrido herdaria dois alelos diferentes, um da mãe e um do pai. Tal prole é dita **heterozigota**. Mendel percebeu que alelos diferentes que estão presentes em um heterozigoto devem coexistir muito embora um seja dominante e o outro recessivo e que cada um desses alelos teria uma chance igual de entrar em um gameta quando o heterozigoto se reproduz. Além disso, ele percebeu que a fertilização aleatória com uma população mista de gametas, metade levando o alelo dominante e metade levando o alelo recessivo, produziria alguns zigotos nos quais ambos os alelos eram recessivos. Assim, ele pôde explicar o reaparecimento da característica recessiva na prole de plantas híbridas.

Mendel usou símbolos para representar os fatores hereditários que ele postulou, uma inovação metodológica. Com símbolos, ele descreveu claramente e concisamente os fenômenos hereditários e analisou matematicamente os resultados dos cruzamentos. Ele fez previsões sobre o resultado de futuros cruzamentos. Embora a prática de usar símbolos para analisar problemas genéticos tenha sido muito refinada desde a época de Mendel, os princípios básicos permanecem os mesmos. Os símbolos representam os genes (ou, mais precisamente, seus alelos) e são manipulados de acordo com as regras da herança que Mendel descobriu. Tais manipulações são a essência da análise genética formal. Como uma introdução a este assunto, consideremos a representação simbólica do cruzamento entre plantas altas e baixas de ervilhas (Fig. 3.2).

As duas variedades puras, alta e baixa, são homozigotas para alelos diferentes de um gene que controla a altura da planta. O alelo para baixa, sendo recessivo, é simbolizado pela letra minúscula *d*; o alelo para alta, sendo dominante,

QUADRO 3.1

Resultados de Cruzamentos Monoíbridos de Mendel

Linhagens Parentais	Prole F ₂	Proporção
Plantas altas × plantas baixas	787 altas, 277 baixas	2,84:1
Sementes lisas × sementes rugosas	5.474 lisas, 1.850 rugosas	2,96:1
Sementes amarelas × sementes verdes	6.022 amarelas, 2.001 verdes	3,01:1
Flores violeta × flores brancas	705 violeta, 224 brancas	3,15:1
Vagens infladas × vagens constrictas	882 infladas, 299 constrictas	2,95:1
Vagens verdes × vagens amarelas	428 verdes, 152 amarelas	2,82:1
Flores axiais × flores terminais	651 axiais, 207 terminais	3,14:1

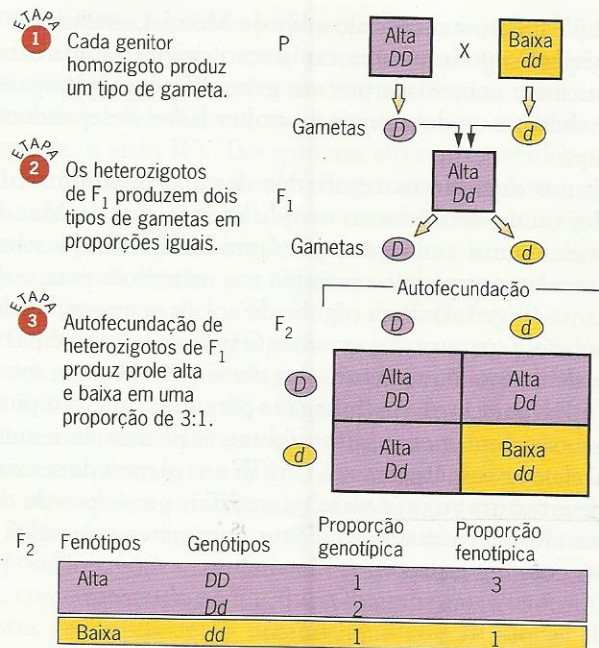


Fig. 3.2 ■ Representação simbólica do cruzamento entre ervilhas altas e baixas e um resumo dos resultados fenotípicos e genotípicos.

é simbolizado pela letra maiúscula correspondente D . Em genética, a letra que é escolhida para indicar os alelos de um gene em geral é tirada da palavra que descreve a característica recessiva (d , de *dwarfness*, anã). Assim, as linhagens alta e baixa são representadas por DD e dd , respectivamente. Diz-se que a constituição alélica de cada linhagem é seu **genótipo**. Em contraste, diz-se que o aspecto físico de cada linhagem, a característica alta ou baixa, é seu **fenótipo**.

Como gerações **parentais**, as plantas de ervilha altas e baixas formam a geração **P** do experimento. Suas proles híbridas são referidas como a primeira geração **filial**, ou F_1 , da palavra latina que significa “filho” ou “filha”. Como cada genitor contribui igualmente para sua prole, o genótipo das plantas F_1 deve ser Dd ; isto é, elas são heterozigotas para os alelos do gene que controla a altura da planta. Seu fenótipo, entretanto, é o mesmo que o da geração parental DD , pois D é dominante em relação a d . Durante a meiose, essas plantas de F_1 produzem dois tipos de gametas, D e d , em iguais proporções. Nenhum dos alelos mudou por ter coexistido com o outro em um genótipo heterozigoto; ao contrário, eles se separam, ou **segregam**, um do outro durante a formação de gametas. Este processo de segregação alélica é talvez a descoberta mais importante que Mendel fez.

Na autofecundação, os dois tipos de gametas produzidos pelos heterozigotos podem unir-se de todos os modos possíveis. Assim, eles produzem quatro tipos de zigotos (escrevemos primeiro a contribuição do ovócito): DD , Dd , dD e dd . Entretanto, devido a dominância, três destes genótipos têm o mesmo fenótipo. Assim, na geração seguinte, referida como F_2 , as plantas são ou altas ou baixas, em uma proporção de 3:1.

Mendel levou sua análise uma etapa adiante. As plantas de F_2 foram autofecundadas para produzir uma F_3 . Todas as plantas baixas de F_2 produzem apenas prole baixa, demonstrando que eram homozigotas para o alelo d , mas a F_2 das plantas altas compreendeu duas categorias. Aproximadamente um terço delas produziu apenas prole alta, enquanto os outros dois terços produziram uma mistura de prole alta e baixa. Mendel concluiu, corretamente, que o terço que era puro homozigoto DD e os dois terços que estavam segregando eram heterozigotos Dd . Estas proporções, $1/3$ e $2/3$, eram exatamente o que sua análise previu porque, entre as plantas altas de F_2 , os genótipos DD e Dd ocorrem em uma proporção de 1:2.

Resumimos a análise de Mendel deste e de outros cruzamentos monóibridos citando os dois princípios básicos que ele descobriu:

- 1. O Princípio da Dominância:** *Em um heterozigoto, um alelo pode mascarar a presença do outro.* Este princípio é relativo ao funcionamento genético. Alguns alelos evidentemente controlam o fenótipo mesmo quando presentes em uma única cópia. Consideramos a explicação fisiológica deste fenômeno em capítulos posteriores.
- 2. O Princípio da Segregação:** *Em um heterozigoto, dois alelos diferentes segregam-se um do outro durante a formação de gametas.* Este princípio é relativo à transmissão genética. Um alelo é fielmente transmitido para a geração seguinte, mesmo se estiver presente com um alelo diferente em um heterozigoto. A base biológica deste fenômeno é o pareamento e a subsequente separação dos cromossomos homólogos durante a meiose, um processo discutido no Cap. 2. Consideraremos os experimentos que levaram a esta teoria cromossômica da hereditariedade no Cap. 5.

Cruzamentos Diíbridos:

O Princípio da Distribuição Independente

Mendel também fez experimentos com plantas que diferiam em duas características (Fig. 3.3). Ele cruzou plantas que produziam sementes amarelas lisas com plantas que produziam sementes verdes rugosas. O objetivo deste experimento era ver se as duas características da semente, cor e textura, eram herdadas independentemente. Como as sementes de F_1 eram todas amarelas e lisas, os alelos para estas duas características eram dominantes. Mendel cultivou plantas destas sementes e as deixou autofecundar-se. Ele então classificou as sementes de F_2 e as contou pelo fenótipo.

As quatro classes fenotípicas na F_2 representam todas as combinações possíveis de cor e textura das características. Duas classes, sementes amarelas e lisas e sementes verdes e rugosas, assemelharam-se às linhagens parentais. As outras duas, sementes verdes e lisas e sementes amarelas e rugosas, mostraram novas combinações de características. As quatro classes tinham uma proporção aproximada de 9 amarelas e lisas:3 verdes e lisas:3 amarelas e rugosas:1 verde e rugosa

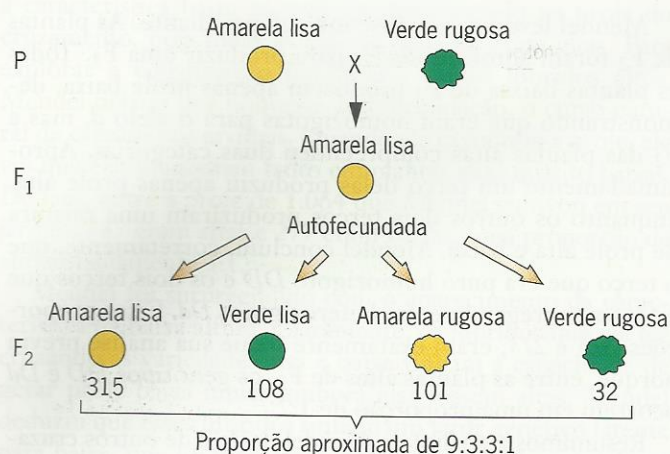


Fig. 3.3 ■ Os cruzamentos de Mendel entre ervilhas que produziram sementes amarelas e lisas e ervilhas que produziram sementes verdes e rugosas.

(Fig. 3.3). Para a mente aguçada de Mendel, estas relações numéricas sugeriram uma explicação simples: Cada característica era controlada por um gene diferente segregando dois alelos, e os dois genes eram herdados independentemente.

Vamos analisar os resultados deste **cruzamento diíbrido**, ou de dois fatores, usando os métodos de Mendel. Representamos cada gene por uma letra, usando a letra minúscula para o alelo recessivo e a maiúscula para o dominante (Fig. 3.4). Para o gene de cor da semente, os dois alelos são *g* (para verde, *green*) e *G* (para amarelo) e para o gene de textura da semente, eles são *w* (para rugoso, *wrinkled*) e *W* (para liso). As linhagens parentais, que são puras, devem ser duplamente homozigotas; as plantas de semente amarela e textura lisa eram *GG WW* e as plantas de semente verde e textura rugosa eram *gg ww*. Tais genótipos de dois genes são habitualmente escritos com pares separados de alelos com um espaço.

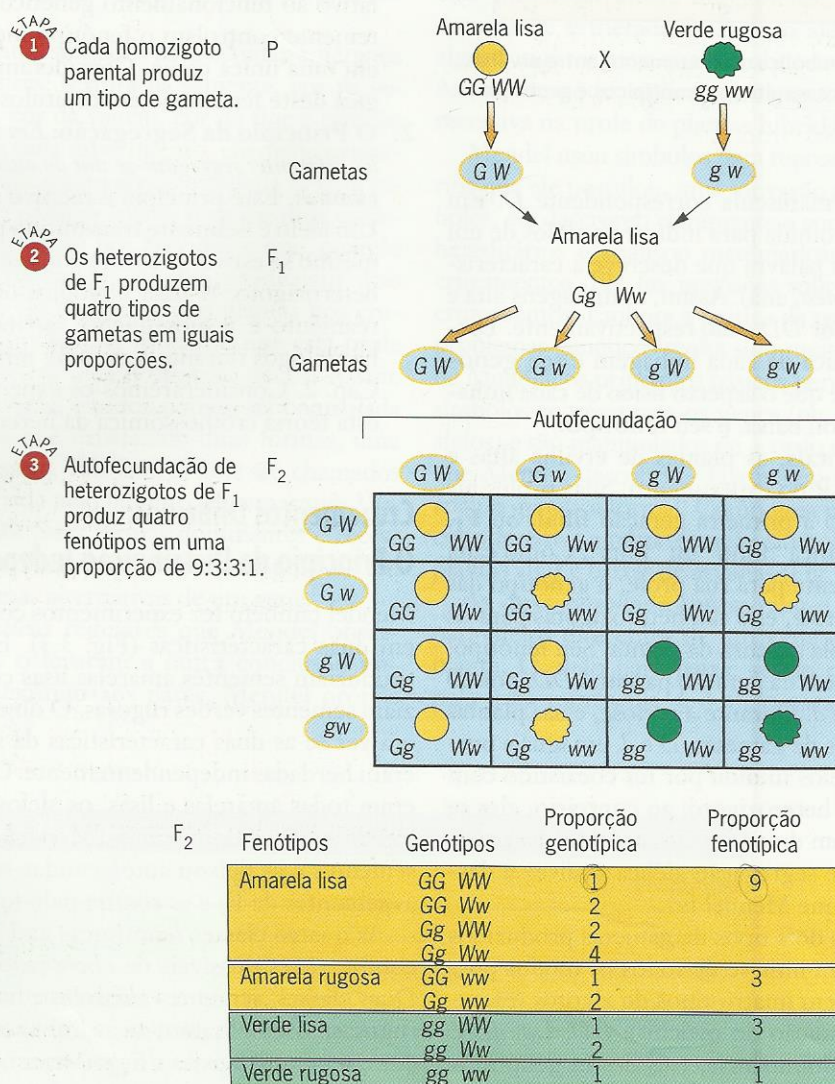


Fig. 3.4 ■ Representação simbólica dos resultados de um cruzamento entre uma variedade de ervilhas com sementes amarelas e lisas e uma variedade com sementes verdes e rugosas.

Os gametas haplóides produzidos por uma planta diplóide contêm uma cópia de cada gene. Os gametas das plantas $GG\ WW$ portanto contêm uma cópia do gene para cor da semente (o alelo G) e uma cópia do gene para textura da semente (o alelo W). Tais gametas são representados por $G\ W$. Por um raciocínio similar, os gametas das plantas $gg\ ww$ são escritos $g\ w$. A fertilização cruzada destes dois tipos de gametas produz híbridos de F_1 que são duplamente heterozigotos, simbolizados por $Gg\ Ww$, e seu fenótipo amarelo e liso indica que os alelos G e W são dominantes.

O Princípio da Segregação prevê que os híbridos de F_1 irão produzir quatro genótipos gaméticos diferentes: (1) GW , (2) Gw , (3) gW e (4) gw . Se cada gene segrega seus alelos independentemente, estes quatro tipos serão igualmente freqüentes; isto é, cada um será 25% do total. Com esta suposição, a autofecundação na F_1 produzirá uma gama de 16 genótipos zigóticos igualmente freqüentes. Obtemos a disposição zigótica combinando sistematicamente os gametas, como mostrado na Fig. 3.4. Obtemos então os fenótipos destes genótipos de F_2 notando que G e W são os alelos dominantes. Em conjunto, existem quatro fenótipos distinguíveis, com freqüências relativas indicadas pelo número de posições ocupadas no arranjo. Para as freqüências absolutas, dividimos cada número pelo total, 16:

amarela, lisa	9/16
amarela, rugosa	3/16
verde, lisa	3/16
verde, rugosa	1/16

Esta análise é prevista por duas suposições: (1) cada gene segrega seus alelos e (2) tais segregações são independentes umas das outras. A segunda suposição significa que não há conexão ou ligação entre os eventos de segregação dos dois genes. Por exemplo, um gameta que recebe W pela segregação do gene de textura tem a mesma possibilidade de receber G que de receber g pela segregação do gene de cor.

Os dados experimentais ajustam-se às previsões de nossa análise? A Fig. 3.5 compara as freqüências previstas e as observadas dos quatro fenótipos de F_2 de dois modos – por proporções e por freqüências numéricas. Para as freqüências numéricas, nós calculamos os números previstos multiplicando a proporção prevista pelo número total de sementes da F_2 examinadas. Para ambos os métodos, obviamente há uma boa concordância entre as observações e as previsões. Assim, as suposições nas quais fazemos nossa análise – segregação independente dos genes da cor da semente e textura da semente – são consistentes com os dados observados.

Mendel fez experimentos similares com outras combinações de características e em cada caso observou que os genes segregavam independentemente. Os resultados desses experimentos o levaram ao terceiro princípio básico:

3. O Princípio da Distribuição Independente: Os alelos de genes diferentes segregam-se, ou como às vezes dizemos, distri-





Fenótipos de F_2	Observada		Esperada	
	Número	Proporção	Número	Proporção
 Amarela lisa	315	0,567	313	0,563
 Verde lisa	108	0,194	104	0,187
 Amarela rugosa	101	0,182	104	0,187
 Verde rugosa	32	0,057	35	0,063
Total	556	1,000	556	1,000

Fig. 3.5 ■ Comparação entre os resultados observados e os esperados na F_2 do experimento de Mendel envolvendo os genes para cor da semente e textura da semente em ervilhas.

buem-se, independentemente uns dos outros. Este princípio é outra regra da transmissão genética, baseado, como veremos no Cap. 5, no comportamento de pares diferentes de cromossomos durante a meiose. Entretanto, nem todos os genes seguem o Princípio da Distribuição Independente. No Cap. 7, consideraremos algumas exceções importantes.

PONTOS IMPORTANTES

- Mendel estudou a herança de sete características diferentes em ervilhas de jardim, cada uma sendo controlada por um gene diferente.
- A pesquisa de Mendel o levou a formular três princípios de herança: (1) os alelos de um gene são dominantes ou recessivos, (2) alelos diferentes de um gene segregam-se uns dos outros durante a formação de gametas e (3) os alelos de genes diferentes distribuem-se independentemente.

APLICAÇÕES DOS PRINCÍPIOS DE MENDEL

Os princípios de Mendel podem ser usados para prever os resultados de cruzamentos entre linhagens diferentes de organismos.

Se a base genética de uma característica é conhecida, os princípios de Mendel podem ser usados para prever o resultado dos cruzamentos. Existem três procedimentos gerais, dois baseados na representação sistemática de todos os genótipos zigóticos ou fenótipos e um baseado na abordagem matemática.

O Método do Quadrado de Punnett

Para situações que envolvem um ou dois genes, é possível escrever todos os gametas e combiná-los sistematicamente para gerar uma gama de genótipos zigóticos. Uma vez que eles tenham sido obtidos, pode ser usado o Princípio da Dominância para determinar os fenótipos associados. Este procedimento, chamado **método do quadrado de Pun-**

nett em homenagem ao geneticista britânico R. C. Punnett, é um modo direto de prever o resultado de cruzamentos. Costumamos usá-lo para analisar o resultado zigótico do cruzamento com os híbridos mendelianos de F₁ amarelo e liso – um tipo de cruzamento comumente denominado **entrecruzamento** (Fig. 3.4). Entretanto, em situações mais complicadas, como as que envolvem mais de dois genes, o método do quadrado de Punnett não é recomendável. Veremos adiante como o método do quadrado de Punnett está relacionado a uma abordagem de problemas genéticos que usam o conceito de probabilidade.

O Método da Linha Bifurcada

Outro procedimento para prever o resultado de um cruzamento envolvendo dois ou mais genes é o **método da linha bifurcada**. Entretanto, em vez de enumerar a prole em um quadrado combinando sistematicamente os gametas, nós a colocamos em um diagrama de linhas ramificadas.

Como exemplo, consideremos um entrecruzamento de ervilhas que são heterozigotas para três genes que se distribuem independentemente, um controlando a altura da planta, um controlando a cor da semente e um controlando

a textura da semente. Este é um cruzamento triíbrido – $Dd Gg Ww \times Dd Gg Ww$ – que pode ser dividido em três cruzamentos monoíbridos – $Dd \times Dd$, $Gg \times Gg$ e $Ww \times Ww$ –, pois todos os genes se distribuem independentemente. Para cada gene, esperamos que os fenótipos apareçam em uma proporção de 3:1. Assim, por exemplo, $Dd \times Dd$ produzirá uma proporção de 3 plantas altas:1 planta baixa. Usando o método da linha bifurcada (Fig. 3.6), podemos combinar estas proporções separadas em uma proporção fenotípica geral para a prole do cruzamento.

Também podemos usar este método para analisar os resultados de um cruzamento entre vários indivíduos heterozigotos e os indivíduos homozigotos. Este tipo de cruzamento é chamado **cruzamento teste**. Por exemplo, se plantas de ervilhas $Dd Gg Ww$ são cruzadas com plantas $dd gg ww$, podemos prever os fenótipos da prole notando que cada um dos três genes nos genitores heterozigotos segrega alelos dominantes e recessivos em uma proporção de 1:1 e que o genitor homozigoto transmite apenas alelos recessivos desses genes. Assim, os genótipos, e finalmente os fenótipos, da prole deste cruzamento dependem de que alelos são transmitidos pelo genitor heterozigoto (Fig. 3.7).

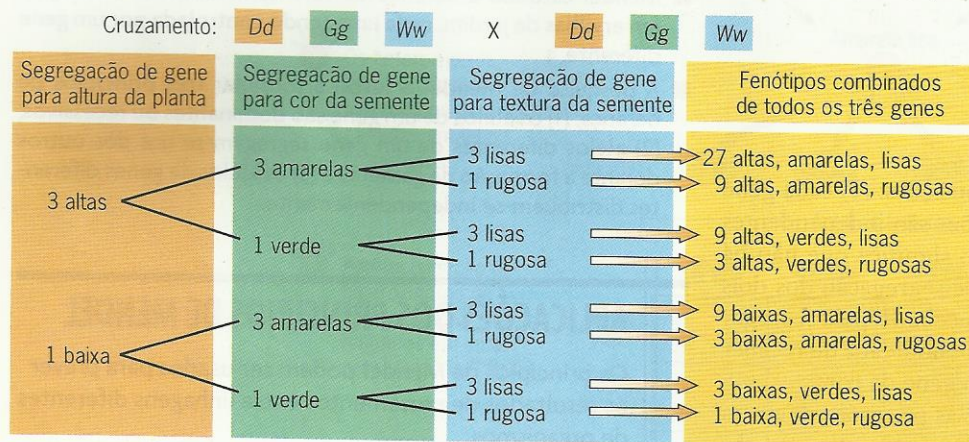


Fig. 3.6 ■ O método da linha bifurcada para prever o resultado de um cruzamento envolvendo três genes que se distribuem independentemente em ervilhas.

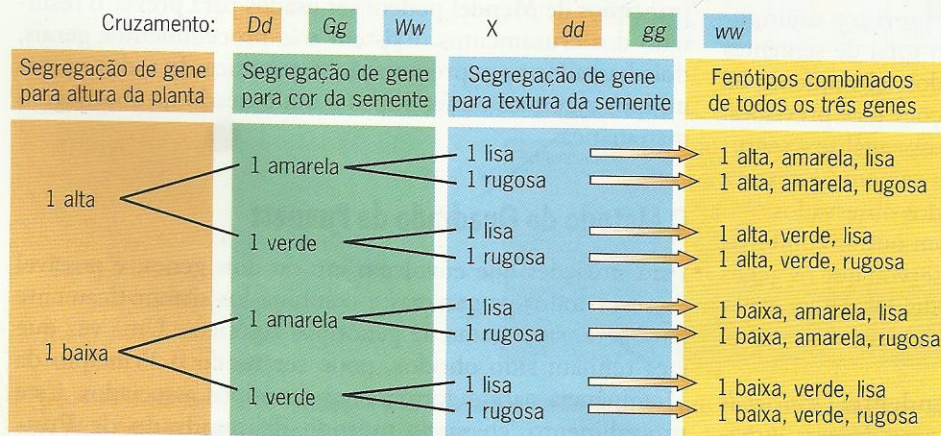


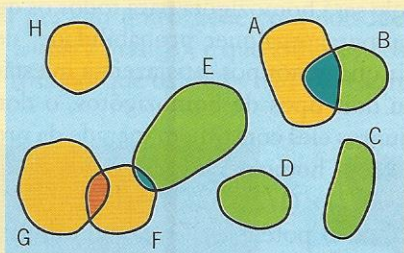
Fig. 3.7 ■ O método da linha bifurcada para prever o resultado de um cruzamento teste envolvendo três genes que se distribuem independentemente em ervilhas.



ENFOQUE TÉCNICO: Regras de Probabilidade

A teoria da probabilidade explica a frequência de eventos – por exemplo, a chance de ter cara ou coroa em um lançamento de moeda, obter um ás de um baralho ou obter um homozigoto dominante de um cruzamento entre dois heterozigotos. Em cada caso, o evento é o resultado de um processo – lançar uma moeda, retirar uma carta, produzir uma prole. Para determinar a probabilidade de um determinado evento, devemos considerar todos os resultados possíveis do processo. A coleção de todos os eventos é chamada *espaço de amostra*. Para o lançamento de uma moeda, o espaço de amostra contém dois eventos, cara e coroa; para retirar uma carta, ele contém 52, um para cada carta; e para os heterozigotos que produzem uma prole, ele contém três, GG, Gg e gg. A probabilidade de um evento é a frequência deste evento no espaço de amostra. Por exemplo, as probabilidades associadas a cada um da prole de um cruzamento entre dois heterozigotos são 1/4 (para GG), 1/2 (para Gg) e 1/4 (para gg).

Em geral, surgem dois tipos de perguntas nos problemas que envolvem probabilidades: (1) Qual a probabilidade de que dois eventos, A e B, ocorram juntos? (2) Qual a probabilidade de que pelo menos um dos dois eventos, A ou B, ocorra? A primeira pergunta especifica a ocorrência conjunta de dois eventos – A e B devem ocorrer juntos para responder a esta pergunta. A segunda é menos restritiva – se A ou B ocorre, a pergunta será respondida. Um diagrama simples pode ajudar a explicar os significados diferentes destas duas perguntas.



As formas no diagrama representam eventos no espaço da amostra, e os tamanhos das formas refletem suas frequências relativas. As superposições entre os espaços indicam a ocorrência conjunta de dois eventos. Se os eventos não se superpõem, eles nunca podem ocorrer juntos. A primeira pergunta é relativa à probabilidade de que A e B ocorram; esta probabilidade é

representada pelo tamanho da superposição entre os dois eventos. A segunda pergunta é relativa à probabilidade de que A ou B ocorra; esta probabilidade é representada pelas formas combinadas dos dois eventos, incluindo, é lógico, a superposição delas.

A Regra Multiplicativa: Se os eventos A e B são independentes, a probabilidade de que eles ocorram juntos, representada por $P(A \text{ e } B)$, é $P(A) \times P(B)$.

Aqui, $P(A)$ e $P(B)$ são as probabilidades dos eventos individuais. Note que independente não significa que eles não sejam superpostos no espaço da amostra. De fato, a não-superposição, ou disjunção, constitui eventos não independentes, pois, se ocorrer um, o outro não ocorrerá. Na teoria das probabilidades, independente significa que um evento não dá informação sobre o outro. Por exemplo, se uma carta retirada de um baralho for um ás, nós não temos informações sobre o conjunto de cartas. Assim, tirar um ás de copas representa a ocorrência conjunta de dois eventos independentes – a carta é um ás (A) e é de copas (H). De acordo com a Regra da Multiplicação, $P(A \text{ e } H) = P(A) \times P(H)$, e como $P(A) = 4/52$ e $P(H) = 1/4$, $P(A \text{ e } H) = (4/52) \times (1/4) = 1/52$.

A Regra da Adição: Se os eventos A e B são independentes, a probabilidade de que pelo menos um deles ocorra, indicado $P(A \text{ ou } B)$, é $P(A) + P(B) - [P(A) \times P(B)]$.

Aqui o termo $P(A) \times P(B)$, que é a probabilidade de que A e B ocorram juntos, é subtraído da soma das probabilidades, $P(A) + P(B)$, pois a soma direta inclui este termo duas vezes. Como exemplo, suponha que procuramos a probabilidade de que uma carta tirada do baralho seja um ás ou copas. De acordo com a Regra da Adição, $P(A \text{ ou } H) = P(A) + P(H) - [P(A) \times P(H)] = (4/52) + (1/4) - [(4/52) \times (1/4)] = 16/52$.

Se os dois eventos não se superpõem no espaço da amostra, a Regra da Adição reduz-se a uma expressão mais simples: $P(A \text{ ou } B) = P(A) + P(B)$. Por exemplo, suponha que procuramos a probabilidade de que uma carta retirada de um baralho seja um ás ou um rei (K). Estes dois eventos não se superpõem no espaço da amostra; eles são ditos mutuamente excludentes. Assim, $P(A \text{ ou } K) = P(A) + P(K) = (4/52) + (4/52) = 8/52$.

O Método da Probabilidade

Um método alternativo ao quadrado de Punnett e o da linha bifurcada, e mais rápido, é baseado no princípio da **probabilidade** (veja Enfoque Técnico: Regras de Probabilidade). A segregação mendeliana é como lançar uma moeda; quando um heterozigoto produz gametas, metade

contém um alelo e metade contém o outro. Se são cruzados dois heterozigotos segregantes, seus gametas combinam-se aleatoriamente, produzindo os genótipos zigóticos (Fig. 3.8). Suponhamos o cruzamento $Aa \times Aa$. A chance de que um zigoto seja AA é simplesmente a probabilidade de que cada um dos gametas que se unem contenha A, ou $(1/2) \times (1/2) = (1/4)$, pois os dois gametas são produzi-

Cruzamento: Aa X Aa

		Gametas masculinos	
		A (1/2)	a (1/2)
Gametas femininos	A (1/2)	AA (1/4)	Aa (1/4)
	a (1/2)	aA (1/4)	aa (1/4)

Prole:

Genótipo	Frequência	Fenótipo	Frequência
AA	1/4	Dominante	3/4
Aa	1/2		
aa	1/4	Recessivo	1/4

Fig. 3.8 ■ Um cruzamento mostrando o método da probabilidade no contexto de um quadrado de Punnett. A frequência de cada genótipo do cruzamento é obtida pelas frequências no quadrado de Punnett, que são, por sua vez, obtidas multiplicando-se as frequências dos dois tipos de gametas produzidos pelos genitores heterozigotos.

dos independentemente. A chance de um homozigoto *aa* é também 1/4. Entretanto, a chance de um heterozigoto *Aa* é 1/2, pois existem dois modos de criar um heterozigoto – *A* pode vir de um ovócito e *a* de um espermatozóide, ou vice-versa. Como cada um destes eventos tem uma chance de um quarto de ocorrer, a probabilidade total de que uma prole seja heterozigota é $(1/4) + (1/4) = (1/2)$. Nós obtemos portanto a seguinte **distribuição de probabilidade** dos genótipos do cruzamento *Aa* × *Aa*:

<i>AA</i>	1/4
<i>Aa</i>	1/2
<i>aa</i>	1/4

Nós concluímos que $(1/4) + (1/2) = (3/4)$ da prole terá o fenótipo dominante e 1/4 terá o recessivo.

Para tal situação simples, o uso do método da probabilidade pode parecer desnecessário. Entretanto, em situações mais complicadas, ele é claramente o enfoque mais prático para prever o resultado dos cruzamentos. Considere, por exemplo, um cruzamento entre plantas heterozigotas para quatro genes diferentes, cada um se distribuindo independentemente. Que fração da prole será homozigota para todos os quatro alelos recessivos? Para responder a esta pergunta, consideramos os genes um de cada vez. Para o primeiro gene, a fração da prole que será de homozigotos recessivos é 1/4, como será para o segundo, o terceiro e o quarto genes. Portanto, pelo Princípio da Distribuição Independente, a fração da prole que será homozigota recessiva quádrupla é $(1/4) \times (1/4) \times (1/4) \times (1/4) = (1/256)$. Certamente, usar o método

Cruzamento: Aa Bb X Aa Bb

		Segregação do gene A	
		A- (3/4)	aa (1/4)
Segregação do gene B	B- (3/4)	A- B- $(3/4) \times (3/4) = 9/16$	aa B- $(1/4) \times (3/4) = 3/16$
	bb (1/4)	A- bb $(3/4) \times (1/4) = 3/16$	aa bb $(1/4) \times (1/4) = 1/16$

Prole:

Genótipo	Frequência	Fenótipo	Frequência
A- B-	9/16	Dominante para ambos os genes	9/16
aa B-	3/16	Recessivo para pelo menos um gene	7/16
A- bb	3/16		
aa bb	1/16		

Fig. 3.9 ■ Aplicação do método de probabilidade a um cruzamento envolvendo dois genes. Neste cruzamento, cada gene segrega fenótipos dominantes e recessivos, com probabilidades de 3/4 e 1/4, respectivamente. Como a segregação ocorre independentemente, as frequências dos fenótipos combinados dentro do quadrado são obtidas multiplicando-se as probabilidades marginais. A frequência da prole que mostra o fenótipo recessivo para pelo menos um dos

da probabilidade é um enfoque melhor que o quadrado de Punnett com 256 entradas!

Consideremos agora uma questão ainda mais difícil. Que fração da prole será homozigota para todos os quatro genes? Antes de computar qualquer probabilidade, devemos primeiro decidir que genótipos satisfazem a questão. Para cada gene existem dois tipos de homozigotos, o dominante e o recessivo, e juntos eles constituem metade da prole. A fração da prole que será homozigota para todos os quatro genes será portanto $(1/2) \times (1/2) \times (1/2) \times (1/2) = (1/16)$.

Para ver todo o potencial do método da probabilidade, precisamos considerar mais uma questão. Suponha que o cruzamento é *Aa Bb* × *Aa Bb* e que desejamos saber que fração da prole apresentará o fenótipo recessivo para pelo menos um gene (Fig. 3.9). Três tipos de genótipos satisfariam esta condição: (1) *A- bb* (o traço representa *A* ou *a*), (2) *aa B-* e (3) *aa bb*. A resposta à pergunta deve ser portanto a soma das probabilidades correspondentes a cada um destes genótipos. A probabilidade para *A- bb* é $(3/4) \times (1/4) = (3/16)$, aquela para *aa B-* é $(1/4) \times (3/4) = (3/16)$ e aquela para *aa bb* é $(1/4) \times (1/4) = (1/16)$. Somando-as, vemos que a resposta é 7/16.

■ PONTOS IMPORTANTES

- O resultado de um cruzamento pode ser previsto pela enumeração sistemática de genótipos usando o quadrado de Punnett.
- Quando mais de dois genes estão envolvidos, o método da linha bifurcada ou da probabilidade é empregado para prever o resultado de um cruzamento.

FORMULANDO E TESTANDO HIPÓTESES GENÉTICAS

O teste do qui-quadrado é um modo simples de avaliar se as previsões de uma hipótese genética concordam com os dados de um experimento.

Uma investigação científica sempre começa com observações de um fenômeno natural. As observações levam a idéias e perguntas sobre o fenômeno, e essas idéias ou perguntas são exploradas mais amplamente conduzindo-se outras observações ou fazendo-se experimentos. Uma idéia científica bem formulada é denominada **hipótese**. Os dados coletados de observações ou experimentação permitem que os cientistas testem hipóteses, isto é, determinem se uma hipótese particular deve ser aceita ou rejeitada.

Como exemplo de teste de hipótese em genética, consideremos a herança de cor da flor em boca-de-leão *Antirrhinum majus*, uma planta de jardim muito popular (Fig. 3.10a). Duas linhagens puras foram obtidas de um estoque, uma com flores bem vermelhas e uma pura com flores brancas. Estas diferenças de cor têm uma base genética? Para responder a esta pergunta, as duas linhagens foram cruzadas (Fig. 3.10b) para produzir híbridos F_1 , todos com flores rosa. Quando os híbridos de F_1 foram inter cruzados, eles produziram três tipos de plantas de F_2 : vermelhas (62), rosa (131) e brancas (57), em números mostrados entre parênteses. Como podemos explicar os dados?

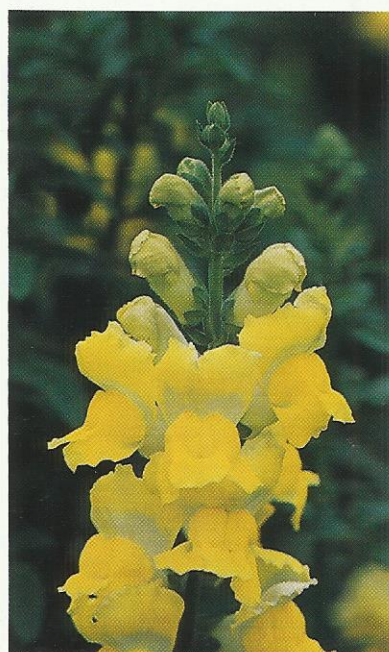
Poderíamos formular uma hipótese de que a cor da flor é controlada por um único gene com dois alelos, W (para vermelho) e w (para branco), e que as flores dos heterozigotos Ww são rosa porque os alelos não são estritamente nem

dominantes nem recessivos. De acordo com esta hipótese, a geração P seria WW (vermelha) $\times ww$ (branca), produzindo híbridos de F_1 que seriam Ww (rosa), que, quando entrecruzados, produziram prole de F_2 WW (vermelha), Ww (rosa) e ww (branca) em uma proporção de 1:2:1. Os números reais parecem apoiar isto, dando crédito a tal hipótese.

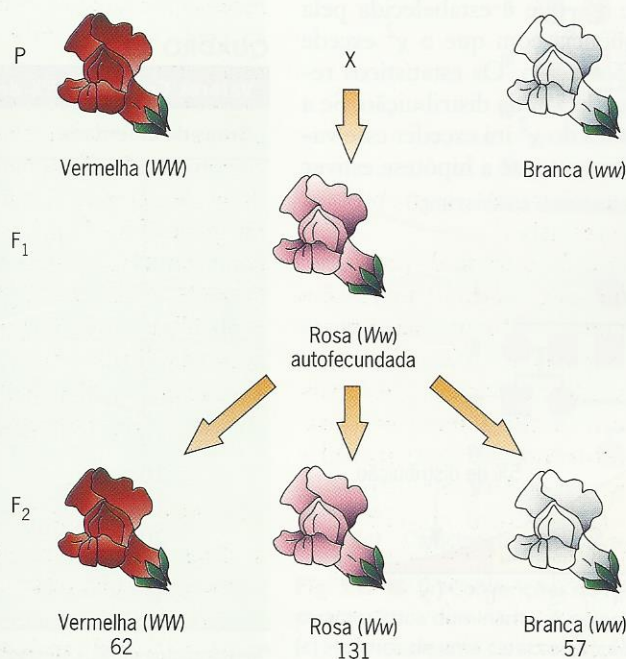
O Teste do Qui-quadrado

Podemos indagar se os dados de fato apóiam uma hipótese particular. Esta pergunta é crítica, pois o valor de uma hipótese depende de sua capacidade de explicar os dados. Uma hipótese que não se ajusta às necessidades precisa ser modificada ou descartada em favor de algo melhor. Um procedimento para teste da correspondência entre a previsão de uma hipótese e os dados reais usa uma estatística chamada de **qui-quadrado** (χ^2). Uma *estatística* é um número calculado a partir de dados – por exemplo, a média de um conjunto de valores examinados. A estatística do χ^2 permite que um pesquisador compare dados, tais como os números que obtemos de um experimento de cruzamento, com seus valores previstos. Se essa comparação for desfavorável, isto é, os dados não corresponderem aos valores previstos, a estatística do χ^2 excederá um número crítico e a hipótese genética será rejeitada. Se a estatística do χ^2 estiver abaixo deste número, a hipótese será aceita. A estatística do χ^2 , portanto, reduz a hipótese testada a um procedimento simples e objetivo.

Como exemplo, consideremos o cruzamento de boca-de-leão descrito antes. Os dados de F_2 parecem consistentes com a hipótese de que um único gene está segregando dois alelos. Entretanto, para avaliar esta hipótese objetivamente, precisamos comparar os dados com os valores previstos. A Fig. 3.11 ilustra os cálculos.






(a)



(b)

Fig. 3.10 ■ (a) Boca-de-leão, *Antirrhinum majus*. (b) Resultados dos cruzamentos entre variedades vermelha e branca de boca-de-leão.

	Fenótipo de F ₂	Número observado	Número esperado
	Vermelha	62	(1/4) x 250 = 62,5
	Rosa	131	(1/2) x 250 = 125
	Branca	57	(1/4) x 250 = 62,5
	Total	250	250

Cálculo da estatística do qui-quadrado para testar a concordância entre números esperados e observados:

$$\chi^2 = \sum \frac{(\text{Observado} - \text{Esperado})^2}{\text{Esperado}}$$

$$= \frac{(62 - 62,5)^2}{62,5} + \frac{(131 - 125)^2}{125} + \frac{(57 - 62,5)^2}{62,5}$$

$$= 0,776$$

Fig. 3.11 ■ Comparação entre os resultados observados e os esperados e o cálculo do χ^2 para um cruzamento de híbridos de boca-de-leão.

O procedimento é direto. Para cada classe fenotípica, computamos o número esperado de prole multiplicando a proporção mendeliana e o tamanho total da amostra. Computamos então a diferença entre os números observados e esperados e fazemos o quadrado destas diferenças para eliminar os efeitos canceladores de valores positivos e negativos. Após dividir o quadrado de cada diferença pelo número esperado correspondente da prole, somamos todos os termos e comparamos a estatística de χ^2 resultante com os valores de distribuição do χ^2 (Fig. 3.12).

A distribuição de valores de χ^2 , que é estabelecida pela teoria estatística, mostra a frequência com que o χ^2 excede um determinado valor apenas por acaso. Os estatísticos recomendam enfocar o valor limite de 5% da distribuição. Se a hipótese estiver correta, a estatística do χ^2 irá exceder esse **valor crítico** de 5% das vezes. Entretanto, se a hipótese estiver

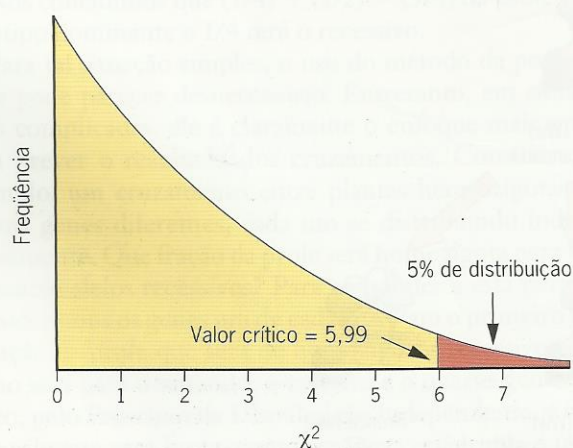


Fig. 3.12 ■ Distribuição de uma estatística de χ^2 .

incorreta, há uma chance muito maior de que o χ^2 irá exceder o valor crítico. Uma hipótese incorreta mais provavelmente produz grandes diferenças entre observações e expectativas. Quando tais grandes diferenças são elevadas ao quadrado, elas produzirão um grande valor da estatística de χ^2 , um situado à direita na escala teórica. É costume rejeitar a hipótese se a estatística do χ^2 exceder o valor crítico. Assim, se a hipótese for verdadeira, há uma chance de 5% de rejeitá-la erradamente.

Voltando ao exemplo: o χ^2 computado, 0,776, deve ser comparado com o valor crítico de uma distribuição teórica. Ocorre, porém, que existem muitas dessas distribuições, e para selecionar a apropriada precisamos conhecer os **graus de liberdade** associados à estatística do χ^2 . Este índice das distribuições de χ^2 é determinado subtraindo-se um do número de classes fenotípicas; neste exemplo, o número de graus de liberdade é $3 - 1 = 2$. Podemos agora comparar a estatística computada de χ^2 com o valor crítico da distribuição teórica com 2 graus de liberdade (veja no Quadro 3.2 uma lista de valores críticos). Devido à estatística computada, 0,776, ser menor que o valor crítico, 5,991, a hipótese de um gene segregando dois alelos *não* é rejeitada. Concluimos que esta hipótese é uma explicação adequada para os dados.

Um problema resolvido ao final do capítulo mostra o que ocorre quando a estatística do χ^2 é maior do que o valor crítico. Outros problemas dão oportunidade para usar a estatística do χ^2 para hipóteses de testes genéticos.

PONTOS IMPORTANTES

- A estatística do qui-quadrado é calculada como $\chi^2 = \sum (\text{número observado} - \text{número esperado})^2 / \text{número esperado}$, com a soma computada para todas as categorias que constituem os dados.
- Cada estatística do qui-quadrado está associada a um índice, os graus de liberdade, que é igual ao número de categorias de dados menos um.

QUADRO 3.2

Valores Críticos^a a 5% de Qui-quadrado (χ^2)

Graus de Liberdade	Valor Crítico a 5%
1	3,841
2	5,991
3	7,815
4	9,488
5	11,070
6	12,592
7	14,067
8	15,507
9	16,919
10	18,307
15	24,996
20	31,410
25	37,652
30	43,773

^aSelecionados de R.A. Fisher e Yates, 1943, *Statistical Tables for Biological, Agricultural, and Medical Research*. Oliver and Boyd, London.

PRINCÍPIOS MENDELIANOS EM GENÉTICA HUMANA

Os princípios mendelianos podem ser aplicados ao estudo da herança de características em humanos.

A aplicação dos princípios mendelianos à genética humana começou logo após a redescoberta da publicação de Mendel em 1900. Entretanto, como não é possível fazer cruzamentos controlados com seres humanos, o progresso foi obviamente lento. A análise genética da hereditariedade humana depende de registros familiares, que são em geral incompletos. Além disso, os seres humanos, ao contrário de organismos experimentais, não produzem muita prole, dificultando discernir as proporções mendelianas, e os humanos não são mantidos e observados em ambiente controlado. O erro de paternidade é outro problema em genética humana, introduzindo um elemento de confusão nos dados. O tempo também é um fator, pois algumas condições genéticas não se manifestam até que uma pessoa atinja a meia-idade. Por todos estes motivos, a análise genética humana tem sido uma tarefa difícil. Entretanto, a vontade de compreender a hereditariedade humana tem sido muito forte, e hoje em dia, a despeito dos obstáculos, aprendemos sobre milhares de genes humanos. O Quadro 3.3 cita algumas das condições que eles controlam. Discu-

QUADRO 3.3

Condições Herdadas em Seres Humanos

Características Dominantes

Acondroplasia (nanismo)
Braquidactilia (dedos curtos)
Cegueira noturna congênita
Síndrome de Ehlers-Danlos (distúrbio do tecido conjuntivo)
Doença de Huntington (distúrbio neurológico)
Síndrome Marfan (alto, estrutura magra)
Neurofibromatose (crescimentos tipo tumorais no corpo)
Sensibilidade gustativa à feniltiocarbamida (PTC)
Bico-de-viúva
Pêlos lanosos

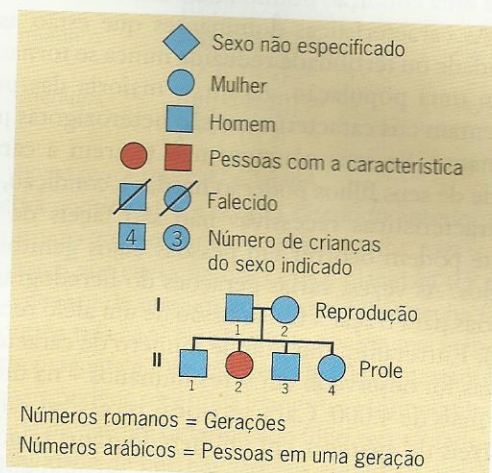
Características Recessivas

Albinismo (falta de pigmento)
Alcaptonúria (distúrbio do metabolismo de aminoácidos)
Ataxia telangiectasia (distúrbio neurológico)
Fibrose cística (distúrbio respiratório)
Distrofia muscular Duchenne
Galactosemia (distúrbio do metabolismo de carboidratos)
Doença de armazenamento de glicogênio
Fenilcetonúria (distúrbio do metabolismo de aminoácidos)
Anemia falciforme (distúrbio da hemoglobina)
Doença Tay-Sachs (distúrbio de estocagem de lipídeos)

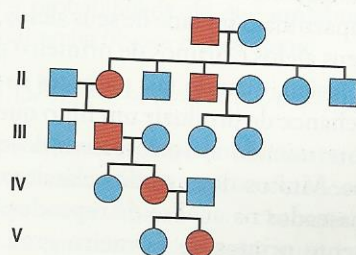
tiremos muitas destas condições em capítulos posteriores deste livro.

Heredogramas

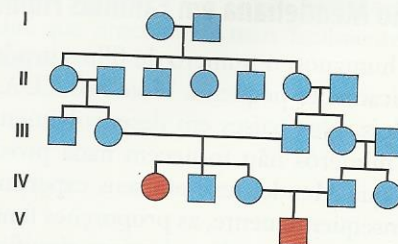
Heredogramas são diagramas que mostram o parentesco entre os membros de uma família (Fig. 3.13a). É costume representar os homens com um quadrado e as mulheres com um círculo. Uma linha horizontal ligando um círculo e um quadrado representa reprodução. A prole da reprodução é mostrada abaixo dos genitores, começando pelo primeiro filho à esquerda e continuando, por ordem de nascimento, para a direita. Pessoas que têm uma condição genética são indicadas



(a) Convenções em um heredograma



(b) Característica dominante



(c) Característica recessiva

Fig. 3.13 ■ (a) Convenções de heredogramas. (b) Herança de uma característica dominante. A característica aparece em cada geração. (c) Herança de uma característica recessiva. As duas pessoas afetadas são prole de parentes.

por cor ou sombreado. As gerações em um heredograma em geral são indicadas em números romanos, e determinadas pessoas dentro de uma geração são referidas por números arábicos depois do número romano.

Características causadas por alelos dominantes são as mais fáceis de identificar. Geralmente, cada pessoa que possui um alelo dominante manifesta a característica, possibilitando rastrear a transmissão do alelo dominante no heredograma (Fig. 3.13b). Cada pessoa afetada deve ter pelo menos um genitor afetado, a menos que, é claro, o alelo dominante tenha aparecido na família como resultado de uma mutação nova – uma mudança em um gene. Entretanto, a frequência da maioria das novas mutações é muito baixa, da ordem de uma em um milhão. Consequentemente, o surgimento espontâneo de uma condição dominante é um evento extremamente raro. As características dominantes que estão associadas a viabilidade ou fertilidade reduzida nunca se tornam frequentes em uma população. Assim, a maioria das pessoas que apresentam tais características são heterozigotas para o alelo dominante. Se seus cônjuges não tiverem a característica, metade de seus filhos poderão herdar a condição.

Características recessivas não são fáceis de identificar porque podem ocorrer em pessoas cujos genitores não são afetados. Às vezes, várias gerações do heredograma são necessárias para traçar a transmissão de um alelo recessivo (Fig. 3.13c). Entretanto, um grande número de características recessivas foi observado em seres humanos – na última contagem, mais de 4.000. Características recessivas raras são mais prováveis de aparecer em um heredograma quando os cônjuges são aparentados – por exemplo, quando são primos em primeiro grau. Esta incidência aumentada ocorre porque os parentes compartilham alelos por terem ancestrais comuns. Irmãos compartilham metade de seus alelos, meio-irmãos um quarto de seus alelos e primos de primeiro grau um oitavo de seus alelos. Assim, quando tais parentes se reproduzem, eles têm maior chance de produzir um filho que seja homozigoto para um determinado alelo recessivo do que genitores não aparentados. Muitos dos estudos clássicos em genética humana são baseados na análise de reprodução entre parentes, principalmente primos de primeiro grau. Consideraremos este assunto em mais detalhes no Cap. 25.

Segregação Mendeliana em Famílias Humanas

Nos seres humanos, o número de filhos produzidos por um casal é tipicamente pequeno. Hoje nos EUA, a média é de cerca de dois. Nos países em desenvolvimento, é de seis a sete. Tais números não fornecem nada próximo do poder estatístico que Mendel teve em seus experimentos com ervilhas. Consequentemente, as proporções fenotípicas em famílias humanas em geral se desviam significativamente de suas expectativas mendelianas.

Como exemplo, consideremos um casal heterozigoto para um alelo recessivo que, em condição homozigota, causa fibrose cística. Se o casal tiver quatro filhos, seria esperado que três não fossem afetados e um fosse afetado por fibrose cística?

A resposta é não. Embora este seja um resultado possível, não é o único. Existem de fato cinco possibilidades distintas:

1. Quatro não afetados, nenhum afetado.
2. Três não afetados, um afetado.
3. Dois não afetados, dois afetados.
4. Um não afetado, três afetados.
5. Nenhum não afetado, quatro afetados.

Intuitivamente, o segundo resultado parece ser o mais provável, pois está de acordo com a proporção mendeliana de 3:1. Podemos calcular a probabilidade deste resultado, e de cada um dos outros, usando os princípios de Mendel e tratando cada nascimento como um evento independente (Fig. 3.14).

Para um determinado nascimento, a chance de que a criança seja não afetada é de $3/4$. A probabilidade de que todos os filhos sejam não afetados é, portanto, $(3/4) \times (3/4) \times (3/4) \times (3/4) = (3/4)^4 = 81/256$. Similarmente, a chance de que um determinado filho seja afetado é de $1/4$; assim, a probabilidade de que todos os quatro sejam afetados é de $(1/4)^4 = 1/256$. Para encontrar as probabilidades dos três outros resultados, precisamos reconhecer que cada um de fato depende de uma coleção de eventos distintos. O resultado de três filhos não afetados e um filho afetado, por exemplo, compreende quatro eventos distintos. Se U representar um filho não afetado e A um filho afetado, e se escrevermos os filhos em sua ordem de nascimento, podemos representar estes eventos como

UUUA, UUAU, UAUU e AUUU

Como cada um tem a probabilidade de $(3/4)^4 \times (1/4)$, a probabilidade total para estes três filhos não afetados e um afetado, independente da ordem de nascimento, é $4 \times (3/4)^3 \times (1/4)$. O coeficiente 4 é o número de modos em que três filhos poderiam não ser afetados e um poderia ser afetado em uma família com quatro filhos. Similarmente, a probabilidade de dois filhos não afetados e dois afetados é $6 \times (3/4)^2 \times (1/4)^2$, pois neste caso existem seis eventos distintos. A probabilidade de um filho não afetado e três afetados é $4 \times (3/4) \times (1/4)^3$, pois neste caso existem quatro eventos distintos. A Fig. 3.14 resume os cálculos sob a forma de probabilidade de distribuição. Como antecipado, três filhos não afetados e um filho afetado é o resultado mais provável (probabilidade de $108/256$). Entretanto, este resultado não é esperado na maioria das vezes devido aos outros quatro resultados terem uma probabilidade combinada de $148/256$. Neste exemplo, a criança cai em duas classes fenotípicas possíveis. Como só existem duas classes, as probabilidades associadas a vários resultados são chamadas **probabilidades binomiais**. O Enfoque Técnico: Probabilidades Binomiais generaliza a análise deste exemplo para outras situações em que os filhos caem em duas classes fenotípicas.

Consulta Genética

O diagnóstico de condições genéticas geralmente é um processo difícil. Tipicamente, os diagnósticos são feitos



ENFOQUE TÉCNICO: Probabilidades Binomiais

A prole de reproduções às vezes segrega em duas classes distintas – por exemplo, homem ou mulher, saudável ou doente, normal ou mutante, fenótipo dominante ou fenótipo recessivo. De modo genérico, podemos referir-nos a estes dois tipos de prole como P e Q e notar que, para qualquer prole, a probabilidade de ser P é p , e a probabilidade de ser Q é q . Como existem apenas duas classes, $q = 1 - p$. Suponha que o número total da prole é n e que cada um é produzido independentemente. Podemos calcular a **probabilidade binomial** de que x da prole cairá em uma classe e y na outra:

Probabilidade de x na classe P e y na classe Q =

$$\frac{n!}{x! y!} p^x q^y$$

O termo entre colchetes contém três funções fatoriais ($n!$, $x!$, e $y!$), cada um dos quais é computado como uma série descendente de produtos. Por exemplo, $n! = n(n-1)(n-2)(n-3)\dots(3)(2)(1)$. Se for necessário $0!$, ele é definido como um. Na fórmula, o termo entre colchetes conta os diferentes modos, ou ordens, em que a prole n pode segregar-se de modo que x caia na classe P e y na classe Q. O outro termo, $p^x q^y$, dá a probabilidade de se obter um modo ou ordem particular. Como cada uma das ordens é igualmente provável, a multiplicação deste termo pelo termo entre colchetes dá a probabilidade de se obter uma prole x na classe P e y na classe Q, independente da ordem de ocorrência.

Se, para valores fixos de n , p e q , variarmos sistematicamente x e y , podemos calcular um conjunto de probabilidades. Este conjunto constitui uma distribuição de probabilidade binomial. Com a distribuição, podemos responder a perguntas tais como “Qual a probabilidade de que x exceda um determinado valor?” ou “Qual a probabilidade de que x fique entre dois determinados valores?” Por exemplo, consideremos uma família com seis filhos. Qual a probabilidade de que pelo menos quatro sejam meninas? Para responder a esta pergunta, nós notamos que, para qualquer filho determinado, a probabilidade de que seja uma menina (p) é $1/2$ e a probabilidade de que seja um menino (q) também é $1/2$. A probabilidade de que exatamente quatro crianças em uma família sejam meninas (e dois sejam meninos) é portanto $[(6!)/(4! 2!)] (1/2)^4 (1/2)^2 = 15/64$, que é um dos termos na distribuição binomial. Entretanto, a probabilidade de que pelo menos quatro sejam meninas (e não mais que dois sejam meninos) é a soma dos três termos para esta distribuição:

Evento	Fórmula Binomial	Probabilidade
4 meninas e 2 meninos	$[(6!)/(4! 2!)] \times (1/2)^4 (1/2)^2 =$	15/64
5 meninas e 1 menino	$[(6!)/(5! 1!)] \times (1/2)^5 (1/2)^1 =$	6/64
6 meninas e 0 menino	$[(6!)/(6! 0!)] \times (1/2)^6 (1/2)^0 =$	1/64

Portanto, a resposta é $(15/64) + (6/64) + (1/64) = 22/64$.

A distribuição binomial também dá respostas a outros tipos de perguntas. Por exemplo, qual a probabilidade de que pelo menos um mas não mais que quatro dos filhos sejam meninas? A resposta é a soma dos quatro termos:

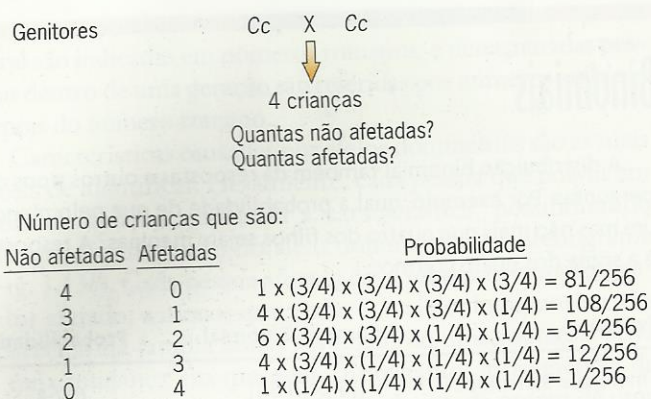
Evento	Fórmula Binomial	Probabilidade
1 menina e 5 meninos	$[(6!)/(1! 5!)] \times (1/2)^1 (1/2)^5 =$	6/64
2 meninas e 4 meninos	$[(6!)/(2! 4!)] \times (1/2)^2 (1/2)^4 =$	15/64
3 meninas e 3 meninos	$[(6!)/(3! 3!)] \times (1/2)^3 (1/2)^3 =$	20/64
4 meninas e 2 meninos	$[(6!)/(4! 2!)] \times (1/2)^4 (1/2)^2 =$	15/64

Resumindo, vemos que a resposta é $56/64$.

Vamos agora considerar o exemplo discutido anteriormente. Um homem e uma mulher, ambos heterozigotos para o alelo mutante recessivo que causa fibrose cística, planejam ter quatro filhos. Qual a chance de que um dos quatro filhos tenha fibrose cística e os outros três não? Nós já vimos por enumeração que a resposta a esta pergunta é $108/256$ (veja Fig. 3.14). Entretanto, esta resposta também poderia ser obtida usando-se a fórmula binomial. A probabilidade de que um determinado filho seja afetado é $p = 1/4$ e a probabilidade de que ele não seja afetado é $q = 3/4$. O número total de filhos é $n = 4$, o número de filhos afetados é $x = 1$ e o número de filhos não afetados é $y = 3$. Juntando tudo, podemos calcular a probabilidade de que exatamente um dos quatro filhos do casal tenha fibrose cística como

$$[4!/(1! 3!)] (1/4)^1 (3/4)^3 = 4 \times (1/4) \times (27/64) = 108/256$$

Nesta expressão, o termo em colchete é o número de modos pelos quais um filho é afetado e três filhos não são afetados em um total de quatro filhos. Este termo é geralmente chamado **coeficiente binomial**. Ele dá o número de modos pelos quais pode ocorrer um evento específico. Para situações complexas envolvendo muitos resultados possíveis, este coeficiente juntamente com outros componentes da fórmula binomial permite-nos calcular as probabilidades que procuramos mais facilmente do que a enumeração exaustiva.



Distribuição de probabilidade

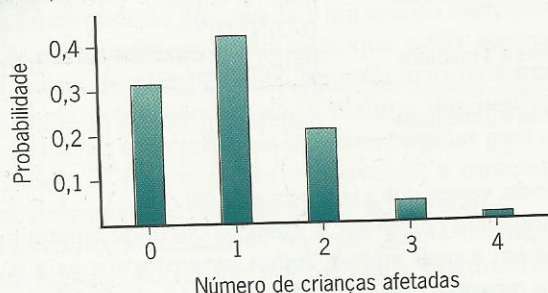


Fig. 3.14 ■ Distribuição de probabilidade de famílias com quatro filhos segregando uma característica recessiva.

por médicos que foram treinados em genética. O estudo destas condições requer grande dose de cuidadosa pesquisa, incluindo exame de pacientes, entrevista com parentes e triagem de estatísticas vitais sobre nascimentos, mortes e casamentos. Os dados acumulados fornecem a base para definir clinicamente a condição e determinar seu modo de herança.

Os prospectivos genitores podem desejar saber se seus filhos correm riscos de herdar determinada condição, especialmente se outros membros da família foram afetados. É responsabilidade do consultor genético avaliar tais riscos e explicá-los aos prospectivos genitores. A avaliação de risco requer familiaridade com probabilidade e estatística, bem como bom conhecimento de genética.

Como exemplo, consideremos um heredograma mostrando a herança do **câncer colorretal não polipóide** (Fig.

3.15). Esta doença é um dos vários tipos de câncer que são herdados. Ele é devido a uma mutação dominante que afeta cerca de 1 em 500 pessoas na população geral. A média de idade em que o câncer colorretal não polipóide aparece em uma pessoa que tem a mutação é de 42 anos. No heredograma, vemos que o câncer se manifestou em pelo menos uma pessoa em cada geração e que todas as pessoas afetadas têm um genitor afetado. Estes fatos são consistentes com o modo dominante de herança desta doença.

A questão da consulta surge na geração V. Entre as nove pessoas mostradas, duas são afetadas e sete não. Cada uma das sete pessoas não afetadas teve um genitor afetado que deve ser heterozigoto para a mutação causadora de câncer. Assim, alguma destas sete pessoas não afetadas pode ter herdado a mutação e teria portanto risco de desenvolver câncer colorretal não polipóide mais tarde na vida. Só o tempo dirá. À medida que as pessoas não afetadas envelhecem, as que possuem a mutação terão um risco aumentado de desenvolver a doença. Assim, quanto mais tempo ficarem não afetadas, maior a probabilidade de que de fato não sejam portadoras. Nesta situação, o risco é uma função da idade da pessoa e deve ser avaliado empiricamente a partir dos dados sobre a idade de início da doença entre pessoas da mesma população, se possível da mesma família. Cada uma das sete pessoas não afetadas terá, é claro, que viver com a ansiedade de ser uma possível portadora da mutação causadora de câncer.* Além disso, em algum ponto terão de decidir se querem reproduzir-se e correr o risco de transmitir a mutação para seus filhos. Discutiremos outros cânceres herdados e os aspectos correlatos da consulta genética no Cap. 24.

Como outro exemplo, considere a situação mostrada na Fig. 3.16. Um casal, indicado por R e S na Fig. 3.16b, está preocupado com a possibilidade de ter um filho (T) com **albinismo**, uma condição autossômica recessiva caracterizada por ausência completa do pigmento melanina na pele, nos olhos e nos cabelos. S, a mãe prospectiva, tem albinismo, e R, o pai prospectivo, tem dois irmãos com albinismo. Parece, portanto, que a criança tem algum risco de nascer com albinismo.

*N.T.: Muitas doenças genéticas já foram mapeadas e podem ser identificadas por testes muito antes de sua manifestação. Você faria o teste para uma doença tardia que ainda não tem tratamento ou cura?

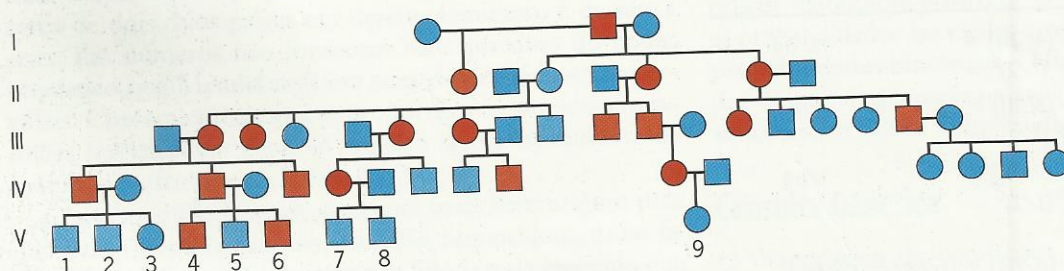


Fig. 3.15 ■ Heredograma mostrando a herança de câncer colorretal não polipóide.



UM MARCO NA GENÉTICA: Publicação de Mendel em 1866

A publicação que iniciou a ciência da genética teve o título “Versuche über Pflanzenhybriden”, que é traduzida do alemão como “Experimentos Com Plantas Híbridas”. Esta publicação detalhou os estudos de Gregor Mendel sobre a herança em ervilhas. Ela foi publicada em 1866 pela Sociedade de História Natural de Brunn.¹ No ano anterior, Mendel havia apresentado os resultados de seus estudos em palestras de duas reuniões da sociedade.

Como a maioria das publicações científicas, a de Mendel continha uma introdução aos aspectos em estudo, uma descrição dos materiais e métodos que ele usou nos experimentos, uma apresentação detalhada dos resultados que ele obteve e uma discussão desses resultados em termos da hipótese subjacente. A publicação é muito objetiva tanto em conteúdo quanto em estilo. Mendel descreveu experimentos com sete características diferentes de ervilhas: textura da semente, cor da semente, cor da flor, forma da vagem, cor da vagem, posição da flor e altura da planta. Seu método foi cruzar variedades contrastantes de ervilhas – por exemplo, plantas altas com plantas baixas – e examinar a prole. Essa prole era então autofecundada, e sua prole, também, era examinada. Mendel continuou a estudar a prole de autofecundação por várias gerações. Os resultados de todo este trabalho deram-lhe a evidência do Princípio da Segregação. Mendel também fez experimentos mais complicados em que duas ou três características eram estudadas simultaneamente em cruzamentos. Os resultados deste trabalho deram-lhe a evidência do Princípio da Distribuição Independente.

Poucas publicações científicas tiveram o impacto que a de Mendel teve. Inicialmente, a publicação foi ignorada. Entretanto, quando seu significado foi reconhecido, tornou-se uma leitura obrigatória para todos os interessados no estudo da hereditariedade. As gerações de leitores levantaram algumas dúvidas sobre a publicação de Mendel. Mendel relatou literalmente seus resultados? A correspondência entre os dados e as previsões de suas hipóteses é muito boa? Ele teria alterado, ou fabricado, os dados para apresentar o que melhor se ajustava a suas hipóteses?

Em 1936, Ronald A. Fisher, um estatístico e geneticista britânico, apresentou uma análise da publicação de Mendel no *Annals of Science*, um periódico devotado à história da ciência.² Fisher tentou cuidadosamente reconstruir o que Mendel fez e quando ele o fez. A pesquisa de Mendel parece ter começado com o cultivo de estoques de ervilhas em 1857. As primeiras hibridizações entre variedades diferentes foram aparentemente feitas em 1858; outras hibridizações foram feitas em 1859. Mendel seguiu a prole desses cruzamentos por seis gerações. Fisher supôs que Mendel começou seus cruzamentos diíbridos e triíbridos em 1861. Nesse ano, ele também aparentemente começou os cruzamentos testes para determinar as proporções gaméticas de plantas heterozigotas. Juntos, os experimentos de Mendel com ervilhas duraram oito anos, de 1857 a 1864. A magnitude desses experimentos é impres-

sionante. Em alguns anos, Mendel cultivou mais de 5.000 plantas de ervilhas no jardim do monastério.

Fisher considera a questão sobre se Mendel relatou ou não os experimentos com ervilhas do seguinte modo:

A publicação de Mendel é, como freqüentemente notado, um modelo com relação a ordem e lucidez com a qual os fatos relevantes foram apresentados, e tal apresentação ordenada teria sido muito facilitada se o autor tivesse a liberdade de ignorar determinados cruzamentos e anos aos quais as plantas que contribuíram para qualquer resultado especial pudessem pertencer. Mendel era um experiente e bem-sucedido professor e poderia muito bem adotar um estilo de apresentação adequado para as palestras sem ter nenhuma obrigação de complicar sua história com detalhes não essenciais. O estilo de apresentação didática, com suas simplificações convencionais, representa, como é bem conhecido, uma tradição muito mais antiga entre os escritores científicos que as narrativas mais literárias nas quais experimentos hoje são habitualmente apresentados.³

Após examinar a evidência na publicação, Fisher conclui que os experimentos de Mendel devem ser considerados literalmente: “Seus experimentos foram feitos do modo e da ordem em que foram relatados.”⁴

Fisher também considera a questão sobre se os dados de Mendel concordam ou não tão bem com as previsões de sua hipótese. Por exemplo, usando os dados que Mendel obteve em 1863, Fisher calculou uma estatística de χ^2 para testar a verossimilhança entre as observações e as expectativas. O resultado, 15,54, é menor do que a metade do valor esperado da distribuição do χ^2 com 41 graus de liberdade, e Fisher disse “que um valor tão baixo raramente ocorre por acaso uma vez em 2.000 tentativas”.⁵ Portanto, ele conclui: “Não pode haver dúvida de que os dados dos últimos anos do experimento foram fortemente tendenciosos na direção da concordância com as expectativas.”⁶

A tendência em favor dos resultados esperados é mais aparente nos experimentos que Mendel fez para determinar se as plantas de F_2 com um fenótipo dominante eram homozigotas ou heterozigotas para o alelo dominante. Seu procedimento foi deixar que as plantas se autofecundassem; então, examinou 10 da prole. Se alguma das 10 mostrasse o fenótipo recessivo, o genitor era classificado como um heterozigoto. Se nenhuma das 10 apresentasse este fenótipo, era classificada como homozigota. Fisher notou que, com este procedimento, alguns heterozigotos seriam incorretamente classificados como homozigotos apenas por acaso. A probabilidade deste erro é $(0,75)^{10} = 0,0563$. Assim, a proporção esperada de genitores segregantes para não segregantes não é 2:1 mas 1,88:1,11. Entre as 600 plantas que Mendel testou, 201 foram classificadas como homozigotas e 399 foram classifi-

¹Mendel, G. 1866. Versuche über Pflanzenhybriden. *Verhandlungen Naturforschender Vereines in Brunn* 10.

²Fisher, R.A. 1936. Has Mendel's work been rediscovered? *Annals of Science* 1:115-137.

³Ibid.

⁴Ibid.

⁵Ibid.

⁶Ibid.



UM MARCO NA GENÉTICA (continuação)

cadadas como heterozigotas. Estes números são muito próximos das expectativas baseadas em uma proporção de 2:1, mas não tão próximos das expectativas baseadas na proporção corrigida para a probabilidade de classificação errada de heterozigotos como homozigotos. Com esta proporção, os números esperados são de 377,5 homozigotos e 222,5 homozigotos. Fisher notou que o desvio dos dados de Mendel destes números previstos “é esperado uma vez em vinte e nove vezes.”⁷ Para esta discrepância, Fisher sugeriu “que Mendel foi iludido por algum assistente que sabia muito bem o que era esperado. Tal possibilidade é apoiada pela evidência independente de que os dados da maioria dos experimentos, se não de todos eles, haviam sido falsificados para concordar bem proximamente com as expectativas de Mendel.”⁸

Em 2004, o respeitado geneticista Edward Novitski publicou um artigo sobre a crítica dos resultados de Mendel por parte de Fisher.⁹ Novitski acredita que, no geral, tais resultados “ajustam-se mais proximamente às proporções ... teoricamente esperadas do que se poderia esperar obter com base no acaso” e que “Particularmente perturbadores são os dois grupos de experimentos nos quais os resultados de Mendel concordam proximamente com as proporções que Mendel pode ter considerado apropriadas mas que foram, de acordo com Fisher (1936), incorretas.”¹⁰ Entretanto, Novitski argumenta que as críticas de Fisher sobre Mendel, especialmente sua sugestão de que alguns dados experimentais podem ter sido falsificados, são infundadas. O procedimento de Mendel para avaliar se uma planta de F₂ com um fenótipo dominante era homozigota ou heterozigota para o alelo dominante pode ter sido afetado por um erro na direção oposta da que Fisher descreveu. Mendel baseou sua decisão sobre o genótipo de uma planta nos fenótipos de 10 de sua prole. Novitski propôs que, para algumas das plantas, Mendel não poderia obter a prole necessária de 10. A taxa de falha na germinação da semente pode ter sido tão alta quanto 2 por cento. Se foi examinada uma prole de menos de 10, e pelo menos uma delas tinha o fenótipo recessivo, então a planta genitora poderia ter sido corretamente classificada como heterozigota. Entretanto, se uma prole menor que 10 fosse examinada – digamos, apenas 8 ou 9 – e nenhuma delas apresentasse o fenótipo recessivo, o que Mendel faria? Novitski conjectura que Mendel não teria classificado a planta parental como homozigota para o alelo dominante. Ele teria excluído essa planta das considerações e a substituído por outra planta que tinha de reserva. A planta excluída seria mais provavelmente uma homozigota dominante e a planta reserva mais provavelmente uma heterozigota. Assim, em um esforço para aderir a esta regra de classificação de uma planta como homozigota dominante apenas após contar uma prole dominante de 10, Mendel pode ter desviado seus resultados em favor de plantas heterozigotas. Este

erro contrabalança o descrito por Fisher e desloca a proporção observada entre heterozigotos e homozigotos dominantes para 2:1, que foi a relatada por Mendel.

Novitski também sugeriu que Mendel pode ter repetido alguns dos experimentos em que os resultados que ele obteve “não pareciam adequar-se a suas expectativas, não com a intenção de iludir, mas para ter certeza de que esses experimentos não eram de fato fidedignos de exceções a suas regras. Tendo obtido dados adicionais, ele pode ter ou substituído os dados desviantes anteriores por números ‘melhores’ ou combinado os dois grupos de dados, os quais geralmente obscurecem a amplitude dos desvios no primeiro grupo ... Pode ser observado que tais procedimentos provavelmente existem no preparo de dados para publicação mesmo hoje em dia.”¹¹

Quanto à possibilidade de que Mendel tenha alterado ou manipulado seus dados, Novitski disse que, se ele o fez, suas alterações “são baseadas não no desejo de enganar, mas como uma concessão a sua audiência despreparada. Podemos imaginar que, na época de escrever para a apresentação oral, Mendel mudou, para fins didáticos, alguns resultados específicos que desviariam sua audiência do tema principal do artigo devido a sua natureza aparentemente aberrante. Certamente palavras tais como fraude ou desonestidade devem ser usadas com cautela. Talvez sua situação possa ser comparada à do competente professor de ciência do colégio que, ao explicar a estrutura do átomo a seus estudantes, cai no modelo simples de Bohr, ciente de que, embora não seja o correto, é o apropriado para a audiência a qual se dirige.”¹²

QUESTÕES PARA DISCUSSÕES

1. Hoje em dia, muitas publicações científicas contêm quatro principais seções de texto – Introdução, Materiais e Métodos, Resultados e Discussão – e são geralmente precedidas por um resumo (ou sumário) e seguidas de uma bibliografia de referências citadas. De que modo este formato varia entre periódicos que publicam artigos sobre genética (por ex., *Cell*, *Genetics*, *Proceedings of the National Academy of Sciences*, *Nature* e *Science*)? De que modo isto difere do formato de publicações de outras disciplinas tais como sociologia, legislação, história e literatura?
2. Gregor Mendel parece ter feito suas pesquisas para satisfazer sua própria curiosidade. Ele não se beneficiou financeira ou profissionalmente da pesquisa, e, exceto o espaço no jardim do monastério e talvez a ajuda de uns poucos assistentes, seu trabalho não precisou de nenhum outro tipo de apoio, digamos, por exemplo, uma verba governamental. O que estas circunstâncias dizem sobre o tipo de pessoa que Mendel era? Como ele pode ser comparado a cientistas profissionais que hoje trabalham em projetos de pesquisa?

⁷Ibid.

⁸Ibid.

⁹Novitski, E. 2004. On Fisher's criticism of Mendel's results with the garden pea. *Genetics* 166:1133-1136.

¹⁰Ibid.

¹¹Ibid.

¹²Ibid.



(a)

Este risco depende de dois fatores: (1) a probabilidade de que R seja um portador heterozigoto do alelo de albinismo (a) e (2) a probabilidade de que ele transmita este alelo para T, se de fato ele for um portador. S, que obviamente é homozigota para o alelo de albinismo, irá transmitir este alelo para sua prole.

Para determinar a primeira probabilidade, precisamos considerar os possíveis genótipos para R. Um deles, de que ele seja homozigoto para o alelo recessivo (aa), está excluído, pois sabemos que ele não é albino. Entretanto, os outros dois genótipos, AA e Aa , permanecem como possibilidades distintas. Para calcular as probabilidades associadas a cada um deles, notamos que ambos os genitores de R teriam de ser heterozigotos, pois eles têm dois filhos com albinismo. A reprodução que produziu R foi portanto $Aa \times Aa$, e de tal reprodução esperaríamos 2/3 da prole sem albinismo como sendo Aa e 1/3 como sendo AA (Fig. 3.16c). Assim, a probabilidade de que R seja um portador heterozigoto do alelo de albinismo é 2/3. Para determinar a probabilidade de que ele transmitirá esse alelo para seu filho, notamos que a estará presente em metade de seus gametas.

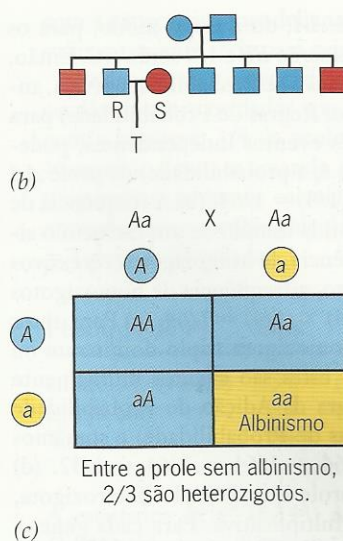


Fig. 3.16 ■ (a) Um filho de índios Cuna do Panamá com albinismo. (b) Heredograma mostrando a herança de albinismo. (c) Quadrado de Punnett mostrando que, entre a prole sem albinismo, a frequência de heterozigotos é 2/3.

Em resumo, o risco de que T seja aa

$$= [\text{Probabilidade de que R seja } Aa] \times [\text{Probabilidade de que R transmita } a, \text{ supondo que R seja } Aa]$$

$$= (2/3) \times (1/2) = (1/3)$$

O exemplo na Fig. 3.16 ilustra uma situação simples de consulta na qual o risco pode ser determinado com precisão. Em geral, as circunstâncias são muito mais complicadas, dificultando a avaliação do risco. A responsabilidade do consultor é analisar a informação do heredograma e determinar o risco o mais precisamente possível.

PONTOS IMPORTANTES

- Heredogramas são usados para identificar características dominantes e recessivas em famílias humanas.
- A análise de heredogramas permite que os consultores genéticos avaliem o risco de uma pessoa herdar determinada característica.

Exercícios Básicos Ilustram a análise genética básica.

1. Duas linhagens de camundongo altamente endocruzadas, uma de pelagem preta e outra de pelagem cinza, foram cruzadas, e toda a prole tinha pelagem preta. Preveja o resultado do entrecruzamento da prole.

Resposta: As duas linhagens de camundongo são evidentemente homozigotas para alelos diferentes de um gene que controla a cor da pelagem: G para preta e g para cinza; o alelo G é dominante porque todos os animais da F_1 são pretos. Quando esses camundongos, genotipicamente Gg , são entrecruzados, os alelos G e g irão segregar um do outro para produzir uma população F_2 que consiste em três genótipos, GG , Gg e gg , na proporção de 1:2:1. Entretanto, devido à dominância do alelo G , os genó-

tipos GG e Gg terão o mesmo fenótipo (pelagem preta); assim, a proporção fenotípica na F_2 será 3 pretos:1 cinza.

2. Uma planta heterozigota para três genes que se distribuem independentemente, $Aa Bb Cc$, é autofecundada. Entre a prole, preveja a frequência de (a) indivíduos $AA BB CC$, (b) indivíduos $aa bb cc$, (c) indivíduos que são ou $AA BB CC$ ou $aa bb cc$, (d) indivíduos $Aa Bb Cc$, (e) indivíduos que não são heterozigotos para todos os três genes.

Resposta: Como os genes se segregam independentemente, podemos analisá-los um de cada vez para obter as respostas a cada uma das perguntas. (a) Quando indivíduos Aa são auto-

fecundados, 1/4 da prole será AA ; do mesmo modo, para os genes B e C , 1/4 dos indivíduos serão BB e 1/4 serão CC . Então, aplicando a Regra da Multiplicação de Probabilidades (veja, anteriormente, Enfoque Técnico: Regras de Probabilidade) para a ocorrência conjunta dos três eventos independentes, podemos calcular a frequência (isto é, a probabilidade) da prole $AA BB CC$ como $(1/4) \times (1/4) \times (1/4) = 1/64$. (b) A frequência de indivíduos $aa bb cc$ pode ser obtida usando-se um raciocínio similar. Para cada gene, a frequência de homozigotos recessivos entre a prole é de 1/4. Portanto, a frequência de homozigotos triplo-recessivos é $(1/4) \times (1/4) \times (1/4) = 1/64$. (c) Para obter a frequência da prole que é homozigota triplo-dominante ou homozigota triplo-recessiva – estes são eventos mutuamente exclusivos –, aplicamos a Regra da Adição de Probabilidade (veja Enfoque Técnico: Regras de Probabilidade) e somamos os resultados de (a) e (b): $1/64 + 1/64 = 2/64 = 1/32$. (d) Para obter a frequência da prole que é triplo-heterozigota, novamente usamos a Regra Multiplicativa. Para cada gene, a frequência de prole heterozigota é 1/2; assim, a frequência de triplo-heterozigotos deve ser $(1/2) \times (1/2) \times (1/2) = 1/8$. (e) A prole que não é heterozigota para todos os três genes ocorre com uma frequência que é de um menos a frequência calculada em (d). Assim, a resposta é $1 - 1/8 = 7/8$.

3. Duas linhagens puras de ervilhas, uma com plantas altas e flores violeta e outra com plantas baixas e flores brancas, foram cruzadas. Todas as plantas de F_1 eram altas e produziram flores violeta. Quando tais plantas foram retrocruzadas com a linhagem parental de plantas baixas e flores brancas, foi obtida a seguinte prole: 53 plantas altas, flores violeta; 48 altas, brancas; 47 baixas, violeta; 52 baixas, brancas. Os genes que controlam o tamanho da planta e a cor da flor segregam independentemente?

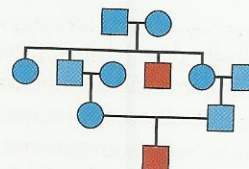
Resposta: A hipótese de distribuição independente do tamanho da planta e cor das flores deve ser avaliada calculando-se um teste estatístico de qui-quadrado dos resultados do experimento. Para obter esta estatística, os resultados devem ser comparados com as previsões da hipótese genética. Na suposição de que os dois genes se distribuem independentemente, as quatro classes fenotípicas na F_2 devem cada uma ser 25% do total (200); isto é, cada uma deve conter 50 indivíduos. Para computar a estatística do qui-quadrado (veja, anteriormente, O Teste do Qui-quadrado), devemos obter a diferença entre cada observação

e seu valor previsto, fazer o quadrado das diferenças, dividir o quadrado de cada diferença pelo valor previsto e então somar os resultados:

$$\chi^2 = (53 - 50)^2/50 + (48 - 50)^2/50 + (47 - 50)^2/50 + (52 - 50)^2/50 = 0,52$$

Esta estatística deve ser comparada com o valor crítico da frequência de distribuição do qui-quadrado para 3 graus de liberdade (calculada como o número de classes fenotípicas menos um). Como o valor computado na estatística do qui-quadrado (0,52) é muito menor que o valor crítico (7,82; veja Quadro 3.2), não há evidência para rejeitar a hipótese de distribuição independente dos genes para tamanho da planta e cor da flor. Assim, podemos hipoteticamente aceitar a idéia de que estes dois genes distribuem-se independentemente.

4. A característica que está segregando no heredograma seguinte é devida a um alelo dominante ou recessivo?



Resposta: Ambos os afetados têm dois genitores não afetados, o que não é consistente com a hipótese de que a característica é devida a um alelo dominante. Assim, a característica parece ser devida a um alelo recessivo.

5. Em uma família com três filhos, qual a probabilidade de que dois sejam meninos e uma seja menina?

Resposta: Para responder a esta pergunta, devemos aplicar a teoria de probabilidades binomiais discutida, anteriormente, no Enfoque Técnico: Probabilidades Binomiais. Para qualquer filho, a probabilidade de que seja um menino é de 1/2, e a probabilidade de que seja uma menina é 1/2. Cada filho é produzido independentemente. Portanto, a probabilidade de dois meninos e uma menina é $(1/2)^3$ vezes o número de modos pelos quais dois meninos e uma menina podem aparecer na ordem de nascimento. Este último termo é o coeficiente binomial $[3!/(2! \times 1!)] = 3$. Assim, a resposta final é $3 \times (1/2)^3 = 3/8$.

Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. A fenilcetonúria, uma doença metabólica em humanos, é causada por um alelo recessivo, k . Se dois heterozigotos portadores do alelo se casam e planejam uma família com cinco filhos: (a) Qual a chance de que todas as crianças não sejam afetadas? (b) Qual a chance de que quatro crianças não sejam afetadas e uma seja afetada pela fenilcetonúria? (c) Qual a chance de que pelo menos três filhos não sejam afetados? (d) Qual a chance de que o primeiro filho seja uma menina não afetada?

Resposta: Antes de responder a cada pergunta, note que de uma reprodução entre dois heterozigotos, a probabilidade de que uma determinada criança seja não afetada é 3/4 e a

probabilidade de que uma determinada criança seja afetada é 1/4. Além disso, para qualquer criança nascida, a chance de que seja um menino é 1/2 e a chance de que seja uma menina é 1/2.

- (a) Para calcular a chance de que todas as cinco crianças sejam não afetadas, use a Regra da Multiplicação. Para cada criança, a chance de que seja não afetada é 3/4, e todas as cinco crianças são independentes. Consequentemente, a probabilidade de cinco crianças não afetadas é $(3/4)^5 = 0,237$. Este é o primeiro termo da distribuição de probabilidade binomial com $p = 3/4$ e $q = 1/4$.

- (b) Para calcular a chance de que quatro crianças serão não afetadas e uma afetada, compute o segundo termo da distribuição binomial:

$$= [5!/(4! 1!)] \times (3/4)^4 \times (1/4)^1 = 5 \times (81/1.024) = 0,399$$

- (c) Para encontrar a probabilidade de que pelo menos três crianças não sejam afetadas, some os primeiros três termos da distribuição binomial:

Evento	Fórmula Binomial	Probabilidade
5 não afetados, 0 afetado	$[(5!)/(5! 0!)] \times (3/4)^5 (1/4)^0 =$	0,237
4 não afetados, 1 afetado	$[(5!)/(4! 1!)] \times (3/4)^4 (1/4)^1 =$	0,399
3 não afetados, 2 afetados	$[(5!)/(3! 2!)] \times (3/4)^3 (1/4)^2 =$	0,264
Total		0,900

- (d) Para determinar a probabilidade de que o primeiro filho seja uma menina não afetada, use a Regra da Multiplicação: $P(\text{filho e filha não afetados}) = P(\text{filho não afetado}) \times P(\text{menina}) = (3/4) \times (1/2) = (3/8)$.

2. Camundongos de populações do tipo selvagem têm pelo cinza-marrom (ou *agouti*), mas, em uma linhagem de laboratório, alguns dos camundongos têm pelagem amarela. Um macho amarelo é cruzado com várias fêmeas agouti. As reproduções produziram uma prole de 40, 22 com pelo agouti e 18 com pelo amarelo. Os animais de F_1 são então entrecruzados para produzir uma F_2 , todos os quais são agouti. Similarmente, os animais amarelos de F_1 são entrecruzados, mas sua prole F_2 segrega em duas classes; 30 são agouti e 54 são amarelos. Os cruzamentos subsequentes entre animais amarelos de F_2 também segregam prole amarela e agouti. Qual a base genética dessas diferenças de cor de pelagem?

Resposta: Notamos que o cruzamento agouti \times agouti só produz animais agouti e que o cruzamento amarelo \times amarelo produz uma mistura de amarelo e agouti. Assim, uma hipótese razoável é que a pelagem amarela seja causada por um alelo dominante, A , e que a pelagem agouti seja causada por um alelo recessivo, a . De acordo com esta hipótese, as fêmeas agouti usadas no cruzamento inicial seriam aa e seu parceiro inicial seria Aa . Supomos que o macho seja heterozigoto porque produziu números aproximadamente iguais de prole agouti e amarela na F_1 . Entre eles, os animais agouti devem ser aa e os animais amarelos Aa . Estas atribuições genóticas surgem dos dados de F_2 , que mostram que os camundongos agouti de F_1 são puros e os amarelos de F_1 segregaram-se. Entretanto, a proporção de segregação de amarelo para agouti (54:30) parece estar desalinhada

com a expectativa mendeliana de 3:1. Essa falta de correlação é suficientemente séria para rejeitar a hipótese?

Podemos usar o procedimento de χ^2 para testar a não-concordância entre os dados e as previsões da hipótese. De acordo com a hipótese, 3/4 da prole F_2 do entrecruzamento amarelo \times amarelo deve ser amarelo e 1/4 deve ser agouti. Usando tais proporções, podemos calcular os números esperados de prole em cada classe e então calcular uma estatística de χ^2 com $2 - 1 = 1$ grau de liberdade.

Fenótipo de F_2	Obs.	Esp.	$(\text{Obs.}-\text{Esp.})^2/\text{Esp.}$
amarelo (AA e Aa)	54	$(3/4) \times 84 = 63$	1,286
agouti (aa)	30	$(1/4) \times 84 = 21$	3,857
Total	84	84	5,143

A estatística de χ^2 (5,143) é muito maior que o valor crítico (3,841) para uma distribuição de χ^2 com 1 grau de liberdade. Consequentemente, nós rejeitamos a hipótese de que as cores de pelagem estão segregando de modo mendeliano 3:1.

O que pode responder pela falha das cores de pelagem em segregar como foi suposto? Obtemos um indício notando que os cruzamentos subsequentes amarelo \times amarelo não estabelecem uma linhagem amarela pura. Isto sugere que os animais amarelos são todos heterozigotos Aa e que os homozigotos AA produzidos pelos cruzamentos entre os heterozigotos não sobrevivem até o estágio adulto. A morte embrionária é, de fato, o motivo dos camundongos amarelos estarem sub-representados nos dados de F_2 . O exame dos úteros de fêmeas grávidas revela que cerca de 1/4 dos embriões morrem. Esses embriões mortos devem ser genotipicamente AA . Assim, uma única cópia do alelo A produz um efeito fenotípico visível (pelagem amarela), mas duas cópias causam a morte. Levando em conta essa mortalidade embrionária, podemos modificar a hipótese e prever que 2/3 da prole nativiva de F_2 deve ser amarela (Aa) e 1/3 deve ser agouti (aa). Podemos então usar o procedimento do χ^2 para testar esta hipótese modificada quanto à consistência dos dados.

Fenótipo de F_2	Obs.	Esp.	$(\text{Obs.}-\text{Esp.})^2/\text{Esp.}$
amarela (Aa)	54	$(2/3) \times 84 = 56$	0,071
agouti (aa)	30	$(1/3) \times 84 = 28$	0,143
Total	84	84	0,214

Esta estatística de χ^2 é menor que o valor crítico de uma distribuição de χ^2 com 1 grau de liberdade. Assim, os dados estão de acordo com as previsões da hipótese modificada.

Questões e Problemas

Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 3.1. Com base nas observações mendelianas, preveja os resultados dos seguintes cruzamentos com ervilhas: (a) uma variedade alta (dominante e homozigota) cruzada com uma variedade baixa; (b) a prole de (a) autofecundada; (c) a prole de (a) cruzada com o genitor alto original; (d) a prole de (a) cruzada com o genitor baixo original.

- 3.2. Mendel cruzou ervilhas que produziram sementes lisas com as que produziam sementes rugosas e autofecundou a prole. Na F_2 , ele observou 5.474 sementes lisas e 1.850 sementes rugosas. Usando as letras W e w para os alelos de textura das sementes, faça um diagrama dos cruzamentos de Mendel, mostrando os genótipos das plantas em cada

geração. Os resultados são consistentes com o Princípio da Segregação?

- 3.3. Um geneticista cruzou camundongos selvagens de cor cinza com camundongos brancos (albinos). Toda a prole era cinza. Essa prole foi entrecruzada para produzir uma F_2 , que consistia em 198 cinza e 72 brancos. Proponha uma hipótese para explicar estes resultados, diagrama os cruzamentos e compare os resultados com as previsões da hipótese.
- 3.4. Uma mulher tem uma anomalia rara das pálpebras chamada ptose, que a impede de abrir completamente os olhos. Sua condição é causada por um alelo dominante, P . O pai da mulher tem ptose, mas sua mãe tem pálpebras normais. Sua avó paterna tem pálpebras normais.
- (a) Quais são os genótipos da mulher, de seu pai e de sua mãe?
- (b) Que proporção dos filhos da mulher terão ptose se ela se casar com um homem de pálpebras normais?
- 3.5. Em pombos, um alelo dominante C causa um padrão xadrez nas penas; seu alelo recessivo c produz um padrão uniforme. A coloração das penas é controlada por um gene que se distribui independentemente; o alelo dominante B produz penas vermelhas e o alelo recessivo b produz penas marrons. As aves de uma variedade pura xadrez vermelha são cruzadas com aves de uma variedade uniforme marrom.
- (a) Preveja o fenótipo da prole.
- (b) Se esta prole for entrecruzada, que fenótipos aparecerão na F_2 e em que proporções?
- 3.6. Em camundongos, o alelo C para pelagem colorida é dominante em relação ao alelo c para pelagem branca e o alelo V para comportamento normal é dominante sobre o alelo v para comportamento valsador, uma forma de descoordenação. Cite os genótipos dos genitores em cada um dos seguintes cruzamentos:
- (a) camundongo colorido normal cruzado com branco normal deu uma prole com 29 coloridos normais e 10 coloridos valsadores;
- (b) camundongo colorido normal cruzado com colorido normal produziu 38 coloridos normais, 15 coloridos valsadores, 11 brancos normais e 4 brancos valsadores;
- (c) camundongo colorido normal cruzado com branco valsador produziu 8 coloridos normais, 7 coloridos valsadores, 9 brancos normais, e 6 brancos valsadores.
- 3.7. Em coelhos, o alelo dominante B causa pelagem preta e o alelo recessivo b causa pelagem marrom; para um gene que se distribui independentemente, o alelo dominante R causa pelagem longa e o alelo recessivo r (de *rex*) causa pelagem curta. Um coelho homozigoto com pelagem longa preta é cruzado com um coelho com pelagem curta marrom, e a prole foi entrecruzada. Na F_2 , qual a proporção de coelhos com pelagem longa preta que será homozigota para ambos os genes?
- 3.8. Em gado *shorthorn*, o genótipo RR causa pelagem vermelha, o genótipo rr causa pelagem branca e o genótipo Rr causa pelagem ruão. Um criador possui vacas e touros com pelagem vermelha, branca e ruão. Que fenótipos devem ser esperados dos seguintes cruzamentos e em que proporções?

- (a) vermelho \times vermelho;
- (b) vermelho \times ruão;
- (c) vermelho \times branco;
- (d) ruão \times ruão.

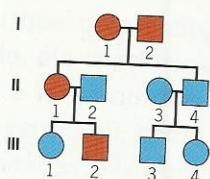
- 3.9. Quantos tipos diferentes de gametas da F_1 , genótipos da F_2 e fenótipos da F_2 seriam esperados dos seguintes cruzamentos:
- (a) $AA \times aa$;
- (b) $AA BB \times aa bb$;
- (c) $AA BB CC \times aa bb cc$
- (d) Que fórmulas gerais são sugeridas por estas respostas?
- 3.10. Um pesquisador estudou seis genes que se distribuem independentemente em uma planta. Cada gene tem um alelo dominante e um recessivo: R caule preto, r caule vermelho; D planta alta, d planta baixa; C vagem cheia, c vagem constricta; O fruto redondo, o fruto oval; H folhas sem penugem, h folhas com penugem; W flor púrpura, w flor branca. Do cruzamento (P_1) $Rr Dd cc Oo Hh Ww \times (P_2)$ $Rr dd Cc oo Hh ww$,
- (a) Quantos tipos de gametas podem ser formados por P_1 ?
- (b) Quantos genótipos são possíveis entre a prole deste cruzamento?
- (c) Quantos fenótipos são possíveis entre a prole?
- (d) Qual a probabilidade de obter o genótipo $Rr Dd cc Oo Hh ww$ na prole?
- (e) Qual a probabilidade de obter um fenótipo preto, baixa, constricta, oval, com penugem, púrpura na prole?
- 3.11. Mendel fez um cruzamento teste de plantas de ervilha cultivadas de sementes amarelas lisas de F_1 com plantas cultivadas de sementes verdes rugosas e obteve os seguintes resultados: 31 amarelas lisas; 26 verdes lisas; 27 amarelas rugosas; e 26 verdes rugosas. Estes resultados são consistentes com a hipótese de que a cor da semente e a textura da semente são controladas por genes que se segregam independentemente, cada um segregando dois alelos?
- 3.12. Para cada uma das situações abaixo, determine os graus de liberdade associados à estatística de χ^2 e decida se o valor observado de χ^2 justifica ou não a aceitação ou rejeição da proporção genética hipotética.

	Proporção Suposta	χ^2 Observado
(a)	3:1	7,0
(b)	1:2:1	7,0
(c)	1:1:1:1	7,0
(d)	9:3:3:1	5,0

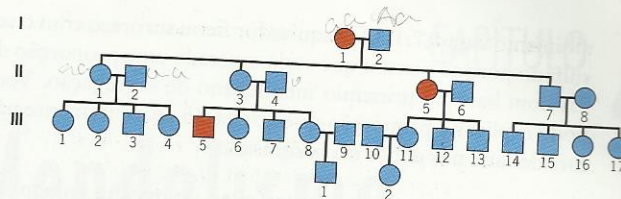
- 3.13. Faça um teste de χ^2 para determinar se uma proporção observada de 30 plantas de ervilha altas:20 plantas de ervilhas baixas é consistente com uma proporção esperada de 1:1 do cruzamento $Dd \times dd$.
- 3.14. As cápsulas das sementes de bolsa-de-pastor são ou triangulares ou ovóides. Um cruzamento entre uma planta com cápsulas de semente triangulares e uma planta com cápsulas de semente ovóides deu híbridos de F_1 todos com cápsulas de semente triangulares. Quando esses híbridos de F_1 foram entrecruzados, eles produziram 80 plantas de F_2 , 72 das quais tinham cápsulas de sementes triangulares e 8 tinham cápsulas de sementes ovóides.

sulas de semente ovóides. Estes resultados são consistentes com a hipótese de que a forma da cápsula é determinada por um único gene com dois alelos?

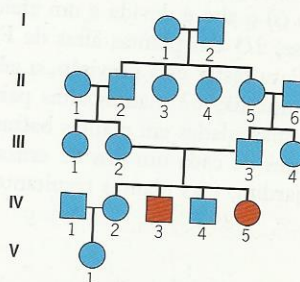
- 3.15. O albinismo em humanos é causado por um alelo recessivo a . Dos casamentos entre pessoas conhecidas como portadoras (Aa) e pessoas com albinismo (aa), que proporção de crianças se esperaria ter albinismo? Entre três crianças, qual a chance de uma sem albinismo e duas com albinismo?
- 3.16. Se tanto o marido quanto a esposa são comprovadamente portadores do alelo para albinismo, qual a chance das seguintes combinações em uma família com quatro filhos: (a) todos os quatro não afetados; (b) três não afetados e um afetado; (c) dois não afetados e dois afetados; (d) um não afetado e três afetados?
- 3.17. Em seres humanos, catarata nos olhos e fragilidade dos ossos são causadas por alelos dominantes que se distribuem independentemente. Um homem com catarata e ossos normais casou com uma mulher sem catarata mas com ossos frágeis. O pai desse homem tinha olhos normais e o pai da mulher tinha ossos normais. Qual a probabilidade de que o primeiro filho deste casal esteja (a) livre de ambas as anomalias; (b) tenha catarata mas não ossos frágeis; (c) tenha ossos frágeis mas não catarata; (d) tenha tanto catarata quanto ossos frágeis?
- 3.18. Na geração V do heredograma da Fig. 3.15, qual a probabilidade de se observar sete crianças sem a mutação causadora de câncer e duas crianças com esta mutação em um total de nove crianças?
- 3.19. Se um homem e uma mulher são heterozigotos para um gene, e se eles tiverem três filhos, qual a chance de que todos os três também sejam heterozigotos?
- 3.20. Se quatro crianças nascem em um determinado dia: (a) Qual a chance de que duas sejam meninos e duas meninas? (b) Qual a chance de que todas as quatro sejam meninas? (c) Que combinação de meninos e meninas entre quatro crianças é mais provável? (d) Qual a chance de que pelo menos uma criança seja uma menina?
- 3.21. Em uma família de seis crianças, qual a chance de que pelo menos três sejam meninas?
- 3.22. O heredograma abaixo mostra a herança de uma característica dominante. Qual a chance de que a prole das seguintes reproduções apresente a característica: (a) III-1 \times III-3; (b) III-2 \times III-4?



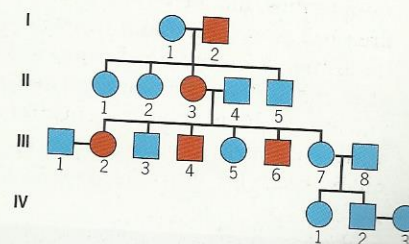
- 3.23. O heredograma a seguir mostra a herança de uma característica recessiva. A menos que exista uma evidência em contrário, suponha que as pessoas que entraram para a família não possuem o alelo recessivo. Qual a chance de que a prole dos seguintes casamentos apresente a característica: (a) III-1 \times III-12; (b) III-4 \times III-14; (c) III-6 \times III-13; (d) IV-1 \times IV-2?



- 3.24. Nos heredogramas abaixo, determine se a característica é mais provavelmente devida a um alelo dominante ou um recessivo. Suponha que esta característica é rara na população.

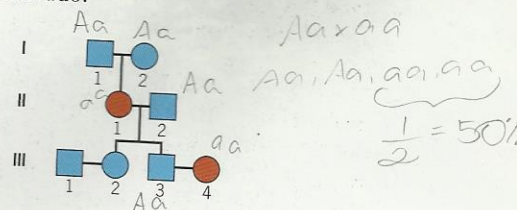


(a)



(b)

- 3.25. No heredograma (b) do Problema 3.24, qual a chance de que o casal III-1 e III-2 tenha um filho afetado? Qual a chance de que o casal IV-2 e IV-3 tenha um filho afetado?
- 3.26. O heredograma abaixo mostra a herança de uma característica recessiva. Qual a chance de que o casal III-3 e III-4 tenha um filho afetado?



- 3.27. Foram entrecruzadas ervilhas heterozigotas para três genes que se distribuem independentemente.
- (a) Que proporção da prole será homozigota para todos os três alelos recessivos?
- (b) Que proporção da prole será homozigota para todos os três genes?
- (c) Que proporção da prole será homozigota para um gene e heterozigota para os outros dois?
- (d) Que proporção da prole será homozigota para o alelo recessivo de pelo menos um gene?

- 3.28. Um pesquisador que esteve estudando albinismo identificou um grande grupo de famílias com quatro filhos em que pelo menos um filho apresentou albinismo. Nenhum dos genitores neste grupo de famílias exibiu albinismo. Entre as crianças, a proporção entre aqueles sem albinismo e aqueles com

albinismo é de 1,7:1. O pesquisador ficou surpreso com o resultado porque pensou que seria esperada uma proporção de 3:1 com base no princípio mendeliano de segregação. Você pode explicar a proporção de segregação aparentemente não mendeliana nos dados do pesquisador?

- 3.29. Um geneticista cruzou plantas altas de ervilha com plantas baixas de ervilha. Todas as plantas de F_1 eram altas. As plantas de F_1 foram autofecundadas, e as plantas F_2 são classificadas por altura: 62 altas e 26 baixas. Destes resultados, o geneticista concluiu que a característica baixa em ervilhas é devida a um alelo recessivo (s) e alta é devida a um alelo dominante (S). Nesta hipótese, $2/3$ das plantas altas de F_2 devem ser heterozigotas Ss . Para testar esta previsão, o geneticista usa pólen de cada uma das 62 plantas altas para fertilizar os óvulos de flores emasculadas em plantas baixas. No ano seguinte, três sementes de cada um dos 62 cruzamentos foram plantadas no jardim, e as plantas resultantes

cresceram até a maturidade. Se nenhuma das três plantas de um cruzamento é baixa, o genitor masculino é classificado como sendo homozigotos SS ; se pelo menos uma das três plantas de um cruzamento é baixa, o genitor masculino é classificado como sendo heterozigoto Ss . Usando este sistema de teste da prole, o geneticista concluiu que 29 das 62 plantas altas de F_2 eram homozigotas SS e 33 destas plantas eram heterozigotas Ss .

- Usando o procedimento do qui-quadrado, avalie estes resultados quanto a verossimilhança com a predição de que $2/3$ das plantas altas de F_2 devem ser heterozigotas.
- Explique por que o procedimento do geneticista para classificar plantas altas de F_2 pelo genótipo não é definitivo.
- Ajuste quanto a incerteza no procedimento de classificação do geneticista e calcule as frequências esperadas de homozigotos e heterozigotos entre as plantas altas de F_2 .
- Avalie as predições obtidas em (c) usando o procedimento do qui-quadrado.

Extensões do Mendelismo

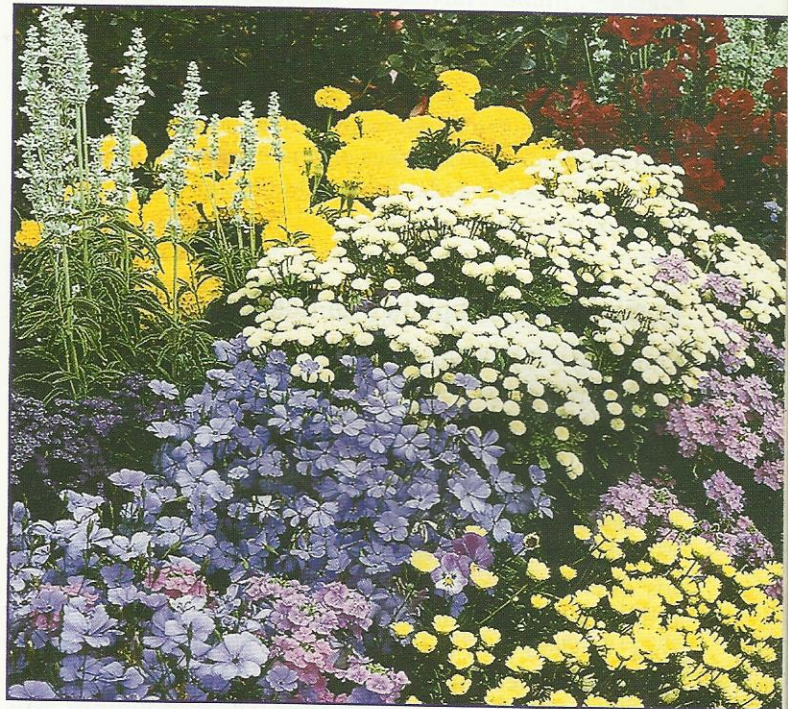
Tópicos

- Variação Alélica e Funcionamento Gênico
- Ação Gênica: Do Genótipo ao Fenótipo

A Genética Cresce Além do Jardim do Monastério de Mendel

Em 1902, entusiasmado pelo que leu no trabalho de Mendel, o biólogo britânico William Bateson publicou uma tradução para o inglês do texto em alemão de Mendel e acrescentou a ele um curto relato sobre o que ele chamou de “Mendelismo — Os Princípios da Dominância, Segregação e Distribuição Independente”. Depois, em 1909, ele publicou o *Mendel's Principles of Heredity*, no qual ele resumiu todas as evidências então disponíveis para apoiar os achados de Mendel. Esse livro foi marcante por dois motivos. Primeiro, ele examinou os resultados de experimentos de cruzamentos com muitas plantas e animais diferentes e em cada caso demonstrou que os princípios de Mendel eram aplicados. Segundo, ele considerou as implicações desses experimentos e levantou dúvidas sobre a natureza fundamental dos genes, ou, como Bateson as chamou, “características unitárias”. Na época em que o livro de Bateson foi publicado, a palavra “gene” ainda não tinha sido inventada.

O livro de Bateson teve um papel crucial em divulgar os princípios do mendelismo no mundo científico. Botânicos, zoólogos, naturalistas, horticulturistas e criadores de animais entenderam a mensagem em



Diversas espécies de plantas crescendo em um jardim. Experimentos com muitas plantas diferentes ampliaram os Princípios de Dominância, Segregação e Distribuição Independente de Mendel.

linguagem simples: os princípios de Mendel, testados por experimentos com ervilhas, feijões, flores, algodão, trigo, cevada, tomates, milho e várias plantas

ornamentais, bem como em gado, ovelhas, gatos, camundongos, coelhos, cobaias, galinhas, pombos, canários e mariposas, eram universais. No prefácio de seu livro, Bateson destacou que “O estudo da hereditariedade tornou-se assim um ramo organizado das ciências fisiológicas, com resultados abundantes e promissórios insuperáveis.”¹

VARIAÇÃO ALÉLICA E FUNCIONAMENTO GÊNICO

Os diversos tipos de alelos dos genes afetam fenótipos de modos diferentes.



Os experimentos de Mendel estabeleceram que os genes existem em formas alternativas. Para cada uma das sete características que ele estudou – cor da semente, textura da semente, altura da

planta, cor da flor, posição da flor, forma da vagem e cor da vagem –, Mendel identificou dois alelos, um dominante e o outro recessivo. Esta descoberta sugeriu uma dicotomia funcional simples entre alelos, como se um alelo não fizesse nada e o outro tudo para determinar o fenótipo. Entretanto, pesquisas do início do século vinte demonstraram que isto é uma simplificação. Os genes podem existir em mais de dois estados alélicos, e cada alelo pode ter um efeito diferente sobre o fenótipo.

Dominância Incompleta e Co-dominância

Um alelo é dominante se tiver o mesmo efeito fenotípico em heterozigotos que em homozigotos, isto é, os genótipos *Aa* e *AA* são fenotipicamente indistinguíveis. As vezes, entretanto, um heterozigoto tem um fenótipo diferente daquele de ambos os homozigotos associados. A cor da flor em boca-de-leão, *Antirrhinum majus*, é um exemplo. As variedades branca e vermelha são homozigotas para alelos diferentes de um gene determinante de cor; quando cruzadas, elas produzem heterozigotos que têm flores rosa. O alelo para cor vermelha (*W*) é portanto considerado **incompletamente**, ou **parcialmente**, **dominante** em relação ao alelo para a cor branca (*w*). A explicação mais provável é que a intensidade de pigmentação nesta espécie depende da quantidade de um produto especificado pelo gene para cor (Fig. 4.1). Se o alelo *W* especifica este produto e o alelo *w* não, os homozigotos *WW* terão o dobro do produto que os heterozigotos *Ww* e portanto apresentarão




Fenótipo	Genótipo	Quantidade do produto gênico
 Vermelha	<i>WW</i>	2x
 Rosa	<i>Ww</i>	x
 Branca	<i>ww</i>	0

Fig. 4.1 ■ A base genética da cor da flor em boca-de-leão. O alelo *W* é incompletamente dominante em relação a *w*. As diferenças entre os fenótipos podem dever-se a diferenças na quantidade do produto especificado pelo alelo *W*.

uma cor mais intensa. Quando o fenótipo do heterozigoto é intermediário aos fenótipos dos dois homozigotos, como aqui, o alelo parcialmente dominante às vezes é dito **semi-dominante** (da palavra latina para “metade”, e portanto semidominante).

Uma outra exceção ao princípio de dominância simples surge quando um heterozigoto apresenta características encontradas em cada um dos homozigotos associados. Isto ocorre com tipos sanguíneos humanos, que são identificados testando-se produtos celulares especiais chamados *antígenos*. Um antígeno é detectado por sua capacidade de reagir com fatores obtidos do soro do sangue. Esses fatores, que são produzidos pelo sistema imunológico, reconhecem antígenos bem especificamente. Assim, por exemplo, um soro, chamado anti-M, reconhece apenas o antígeno M nas células sanguíneas humanas; outro soro, chamado anti-N, reconhece apenas o antígeno N nestas células (Fig. 4.2). Quando um destes soros detecta seu antígeno específico em uma tipagem sanguínea, as células aglomeram-se em uma reação chamada *aglutinação*. Assim, testando-se células para aglutinação com soros diferentes, um técnico de laboratório pode identificar que antígenos estão presentes e portanto determinar o tipo sanguíneo.

A capacidade de produzir os antígenos M e N é determinada por um gene com dois alelos. Um alelo permite que o antígeno M seja produzido; o outro permite que o antígeno N seja produzido. Os homozigotos para o alelo M produ-

¹Bateson, W. 1909. *Mendel's Principles of Heredity*. University Press, Cambridge, Inglaterra.


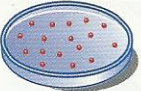


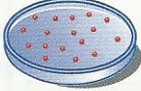

Genótipo	Tipo sanguíneo (antígeno presente)	Reações com anti-soros	
		Soro anti-M	Soro anti-N
$L^M L^M$	M (M)		
$L^M L^N$	MN (M e N)		
$L^N L^N$	N (N)		

Fig. 4.2 ■ A detecção dos antígenos M e N em células sanguíneas por aglutinação com anti-soros específicos. Com soros anti-M e anti-N, podem ser identificados três tipos de sangue.

zem apenas o antígeno M, e os homozigotos para o alelo N produzem apenas o antígeno N. Entretanto, os heterozigotos para estes dois alelos produzem ambos os tipos de antígenos. Como os dois alelos parecem contribuir independentemente para o fenótipo dos heterozigotos, eles são chamados **co-dominantes**. A co-dominância significa que há uma independência de função dos alelos. Nenhum dos alelos é dominante, ou mesmo parcialmente dominante, em relação ao outro. Seria, portanto, impróprio distinguir os alelos por letras maiúsculas e minúsculas, como tivemos em todos os exemplos anteriores. Em vez disso, os alelos co-dominantes são representados por expoentes do símbolo do gene, que neste caso é a letra *L*, um tributo a Karl Landsteiner, o descobridor da tipagem sanguínea. Assim, o alelo M é L^M e o alelo N é L^N . A Fig. 4.2 mostra os três genótipos possíveis formados pelos alelos L^M e L^N e seus fenótipos associados.

Alelos Múltiplos

O conceito mendeliano de que os genes existem em não mais que dois estados alélicos teve que ser modificado quando os genes com três, quatro ou mais alelos foram descobertos. Um exemplo clássico de um gene com **alelos múltiplos** é o que controla a cor da pelagem em coelhos (Fig. 4.3). O gene determinante da cor, indicado pela letra minúscula *c*, tem quatro alelos, três dos quais são distinguidos por um expoente: *c* (*albino*), c^h (*himalaio*), c^{ch} (*chinchilla*) e c^+ (tipo selvagem). Na condição homozigota, cada alelo tem um efeito característico sobre a cor da pelagem: *cc* – pêlos brancos por todo o corpo; $c^h c^h$ – pêlos pretos nas extremidades, pêlos brancos em todo o restante; $c^{ch} c^{ch}$ – pêlos brancos com extremidades do corpo pretas; $c^+ c^+$ – pêlos coloridos por todo o corpo. Como a maioria dos coelhos nas populações selvagens são homozigotos para o alelo c^+ , este alelo é chamado **tipo selvagem**. Em genética, é costume representar os alelos tipo selvagem por um expoente mais após a letra do



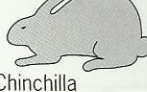
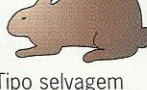
	Genótipo	Fenótipo
	<i>cc</i>	Pêlos brancos por todo o corpo
Albino		
	$c^h c^h$	Pêlos pretos nas extremidades; pêlos brancos no restante
Himalaio		
	$c^{ch} c^{ch}$	Pêlos brancos com pontas pretas no corpo
Chinchilla		
	$c^+ c^+$	Pêlos coloridos por todo o corpo
Tipo selvagem		

Fig. 4.3 ■ Cor da pelagem em coelhos. Os fenótipos diferentes são causados por quatro alelos diferentes do gene *c*.

gene. Quando o contexto é claro, a letra às vezes é omitida, e é usado apenas o sinal mais; assim, c^+ pode ser abreviado simplesmente por um +.

Os outros alelos para o gene *c* são **mutantes** – formas alteradas do alelo tipo selvagem que devem surgir às vezes durante a evolução do coelho. Os alelos *himalaio* e *chinchilla* são indicados por expoentes, mas o alelo *albino* é indicado simplesmente pela letra *c* (para incolor, outro termo para a condição albina). Esta notação reflete outro hábito na nomenclatura genética: os genes são geralmente denominados por um alelo mutante, em geral o alelo associado ao fenótipo mais anormal. A convenção de denominar um gene pelo alelo mutante é em geral consistente com a convenção que discutimos no Cap. 3 – denominar os genes pelo alelo recessivo –, pois a maioria dos alelos mutantes são recessivos. Entretanto, às vezes um alelo mutante é dominante, em cujo caso o gene é denominado por seu fenótipo associado. Por exemplo, um gene em camundongo controla o tamanho da cauda. O primeiro alelo mutante deste gene que foi descoberto causa um encurtamento da cauda em heterozigotos. Este mutante dominante foi portanto simbolizado por *T*, para tamanho da cauda. Todos os outros alelos deste gene, e existem muitos, são indicados por uma letra maiúscula ou minúscula, dependendo de serem dominantes ou recessivos; alelos diferentes são distinguidos uns dos outros por expoentes.

Um outro exemplo de alelos múltiplos vem do estudo de tipos sanguíneos humanos. Os tipos sanguíneos A, B, AB e O, como os tipos sanguíneos M, N e MN já discutidos, são identificados testando-se uma amostra de sangue com soros diferentes. Um soro detecta o antígeno A; outro, o antígeno B. Quando apenas o antígeno A está presente nas células, o tipo sanguíneo é A; quando apenas o antígeno B está presente, o sangue é tipo B. Quando ambos os antígenos estão presentes, o sangue é tipo AB e, quando nenhum antígeno

QUADRO 4.1

Genótipos, Fenótipos e Frequências no Sistema Sangüíneo ABO				
Genótipo	Tipo Sangüíneo	Antígeno A Presente	Antígeno B Presente	Frequência na População Branca dos EUA (%)
$I^A I^A$ ou $I^A i$	A	+	-	41
$I^B I^B$ ou $I^B i$	B	-	+	11
$I^A I^B$	AB	+	+	4
ii	O	-	-	44

está presente, o tipo é O. O tipo sangüíneo para os antígenos A e B é totalmente independente dos tipos de antígenos M e N.

O gene responsável por produzir os antígenos A e B é indicado pela letra *I*. Ele tem três alelos, I^A , I^B e *i*. O alelo I^A especifica a produção do antígeno A e o alelo I^B especifica a produção do antígeno B. Entretanto, o alelo *i* não especifica um antígeno. Entre os seis genótipos possíveis, existem quatro fenótipos distinguíveis – os tipos sangüíneos A, B, AB e O (Quadro 4.1). Neste sistema, os alelos I^A e I^B são co-dominantes, pois cada um é expresso igualmente nos heterozigotos $I^A I^B$, e o alelo *i* é recessivo tanto em relação ao alelo I^A quanto ao I^B . Todos os três alelos são encontrados em frequências apreciáveis em populações humanas; assim, o gene *I* é dito **polimórfico**, para a palavra grega que significa “tendo muitas formas”. Consideraremos o significado populacional e evolutivo dos polimorfismos genéticos no Cap. 27.

Série Alélica

As correlações funcionais entre os membros de uma série de alelos múltiplos podem ser estudadas fazendo-se combinações heterozigotas por meio de cruzamentos entre homozigotos. Por exemplo, os quatro alelos do gene *c* em coelhos podem ser combinados uns com os outros para fazer seis tipos diferentes de heterozigotos: $c^b c$, $c^{ch} c$, $c^+ c$, $c^{ch} c^b$, $c^+ c^b$ e $c^+ c^{ch}$. Estes heterozigotos permitem que sejam estudadas as relações de dominância entre os alelos (Fig. 4.4). O alelo tipo selvagem é totalmente dominante em relação aos outros alelos na série; o alelo *chinchilla* é parcialmente dominante em relação aos alelos *himalaio* e *albino* e o alelo *himalaio* é completamente dominante em relação ao alelo *albino*. Tais relações de dominância podem ser resumidas como $c^+ > c^{ch} > c^b > c$.

Note que a hierarquia de dominância é paralela aos efeitos que os alelos têm sobre a cor da pelagem. Uma explicação plausível é que o gene *c* controla uma etapa na formação de pigmento preto na pelagem. O alelo tipo selvagem é totalmente funcional neste processo, produzindo pêlos coloridos distribuídos no corpo. Os alelos *chinchilla* e *himalaio* são apenas parcialmente funcionais, produzindo alguns pêlos coloridos, e o alelo *albino* não é funcional. Alelos não



Fig. 4.4 ■ Fenótipos de diferentes combinações de alelos *c* em coelhos. Os alelos formam uma série, com o alelo tipo selvagem, c^+ , dominante em relação a todos os outros alelos e o alelo nulo, *c* (*albino*), recessivo em relação a todos os outros alelos; um alelo hipomorfo, c^{ch} (*chinchilla*), é parcialmente dominante em relação a outro, c^h (*himalaio*).

funcionais são ditos **nulos** ou **amorfos** (das palavras gregas para “sem forma”); eles são quase sempre totalmente recessivos. Alelos parcialmente funcionais são ditos **hipomorfos** (das palavras gregas para “com menos forma”); eles são recessivos em relação aos alelos que são mais funcionais, incluindo (geralmente) o alelo tipo selvagem. Mais adiante neste capítulo, consideraremos a base bioquímica para tais diferenças.

Testando Mutações Genéticas Quanto a Alelismo

Um alelo mutante é criado quando um alelo existente muda para um novo estado genético, um processo chamado **mutação**. Este evento sempre envolve uma mudança na composição física do gene (veja Cap. 13) e às vezes produz um alelo que tem um efeito fenotípico detectável. Se, por exemplo, o alelo c^+ sofre mutação para um alelo nulo, um coelho homozigoto para esta mutação teria o fenótipo albino. Entretanto, nem sempre é possível atribuir uma nova mutação a um gene com base em seu efeito fenotípico. Em coelhos, por exemplo, vários genes determinam a cor da pelagem, e uma mutação em qualquer um deles pode reduzir, alterar ou abolir a pigmentação dos pêlos. Assim, se surge uma nova cor em uma população de coelhos, não é imediatamente claro qual gene sofreu mutação.

Um simples teste pode ser usado para determinar a identidade alélica de uma nova mutação, desde que a nova mutação seja recessiva. O procedimento envolve cruzamentos para combinar as novas mutações recessivas com mutações recessivas de genes conhecidos (Fig. 4.5). Se a prole híbrida apresenta um fenótipo mutante, então a nova mutação e a mutação testadora são alelos do mesmo gene. Se a prole híbrida apresentar um fenótipo selvagem, então a nova mutação e a mutação testadora não são alelos do mesmo gene. Este

Mutação recessiva nova	Genótipo testador	Fenótipo híbrido	Conclusão
c^+c^+ X	$a a \Rightarrow$	Tipo selvagem	a e c^+ não alelos
	$b b \Rightarrow$	Tipo selvagem	b e c^+ não alelos
	$c c \Rightarrow$	Mutante	c e c^+ alelos
	$d d \Rightarrow$	Tipo selvagem	d e c^+ não alelos

Fig. 4.5 ■ Um esquema geral destinado a testar mutações recessivas para alelismo. Duas mutações são alelos se um híbrido que contém ambas têm o fenótipo mutante.

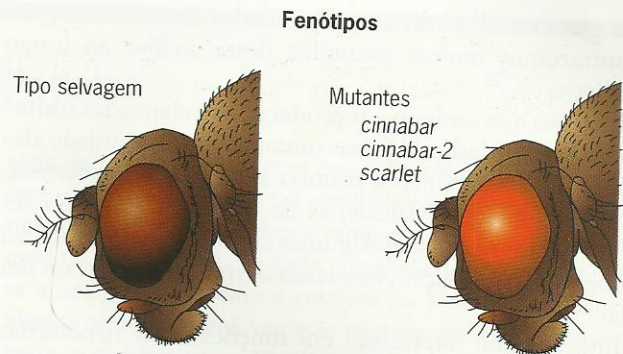
teste é baseado no princípio de que mutações do mesmo gene prejudicam a mesma função genética. Se duas destas mutações são combinadas, o organismo será anormal para esta função e apresentará um fenótipo mutante, mesmo se as duas mutações tiverem uma origem independente.

É importante lembrar que este teste aplica-se apenas a mutações recessivas. Mutações dominantes não podem ser testadas deste modo porque exercem seus efeitos mesmo que esteja presente uma cópia tipo selvagem do gene.

Como exemplo, consideremos a análise de duas mutações recessivas que afetam a cor dos olhos na mosca-das-frutas *Drosophila melanogaster* (Fig. 4.6). Este organismo foi investigado por geneticistas durante quase um século, e muitas mutações diferentes foram identificadas. Duas mutações isoladas independentemente, chamadas *cinnabar* e *scarlet*, são fenotipicamente indistinguíveis, cada uma determinando olhos vermelho-claros. Em moscas tipo selvagem, os olhos são vermelho-escuros. Queremos saber se as mutações *cinnabar* e *scarlet* são alelos de um único gene determinante de cor ou se são mutações em dois genes diferentes. Para encontrar a resposta, devemos cruzar as linhagens mutantes homozigotas entre si para produzir uma prole híbrida. Se os híbridos tiverem olhos vermelho-claros, concluiremos que *cinnabar* e *scarlet* são alelos do mesmo gene. Se tiverem olhos vermelho-escuros, concluiremos que são mutações em genes diferentes.

A prole híbrida tinha olhos vermelho-escuros; isto é, são do tipo selvagem, e não mutante. Assim, *cinnabar* e *scarlet* não são alelos do mesmo gene, mas sim mutações em dois genes diferentes, cada um aparentemente envolvido no controle da pigmentação do olho. Quando testamos uma terceira mutação, chamada *cinnabar-2*, quanto a alelismo com as mutações *cinnabar* e *scarlet*, observamos que a combinação híbrida de *cinnabar-2* e *cinnabar* tem o fenótipo mutante (olhos vermelho-claros) e a combinação híbrida de *cinnabar-2* e *scarlet* tem o fenótipo selvagem (olhos vermelho-escuros). Estes resultados nos dizem que as mutações *cinnabar* e *cinnabar-2* são alelos de um gene determinante de cor e que a mutação *scarlet* não é um alelo deste gene. A mutação *scarlet* define um outro gene determinante de cor.

O teste para determinar se as mutações são alelos de um determinado gene é baseado no efeito fenotípico de combinação das mutações no mesmo indivíduo. Se a combinação híbrida é mutante, nós concluímos que as mutações são alélicas; se for tipo selvagem, concluímos que elas não são



Cruzamento		Cor de olho do híbrido	Conclusão
<i>cinnabar</i>	X <i>scarlet</i>	Tipo selvagem	<i>cinnabar</i> e <i>scarlet</i> são mutações em genes diferentes.
<i>cinnabar-2</i>	X <i>cinnabar</i>	Mutante	<i>cinnabar-2</i> e <i>cinnabar</i> são alelos do mesmo gene.
<i>cinnabar-2</i>	X <i>scarlet</i>	Tipo selvagem	<i>cinnabar-2</i> e <i>scarlet</i> são mutações em genes diferentes.

Fig. 4.6 ■ Um teste para alelismo envolvendo mutações recessivas para cor de olho em *Drosophila*. Três mutações fenotipicamente idênticas, *cinnabar*, *scarlet* e *cinnabar-2*, são testadas para alelismo fazendo-se cruzamentos pareados entre moscas homozigotas para mutações diferentes. Os fenótipos dos híbridos mostram que as mutações *cinnabar* e *cinnabar-2* são alelos de um único gene e que a mutação *scarlet* não é um alelo deste gene.

alélicas. O Cap. 14 discute como este teste, chamado *teste de complementação* na terminologia moderna, permite que geneticistas definam as funções de genes individuais.

Varição Entre os Efeitos de Mutações

Os genes são identificados por mutações que alteram o fenótipo de algum modo claro. Por exemplo, uma mutação pode alterar a cor ou a forma dos olhos, alterar o comportamento ou causar esterilidade e até mesmo a morte. As enormes variações entre os efeitos de mutações individuais sugerem que cada organismo leva muitos tipos diferentes de genes e que cada uma delas pode sofrer mutação de modos diferentes. Na natureza, as mutações fornecem a matéria-prima para a evolução (veja Cap. 27).

Mutações que alteram algum aspecto da morfologia, tal como textura da semente ou cor, são chamadas **mutações visíveis**. A maioria das mutações visíveis são recessivas, mas algumas são dominantes. Os geneticistas aprenderam muito

sobre genes analisando as propriedades dessas mutações. Encontraremos muitos exemplos desta análise ao longo deste livro.

Mutações que limitam a reprodução são chamadas **mutações de esterilidade**. Algumas mutações de esterilidade afetam ambos os sexos, mas a maioria afeta machos *ou* fêmeas. Como nas mutações visíveis, as de esterilidade podem ser dominantes ou recessivas. Algumas de esterilidade impedem totalmente a reprodução, enquanto outras só prejudicam um pouco.

Mutações que interferem em funções vitais necessárias são chamadas **mutações letais**. Seu efeito fenotípico é a morte. Nós sabemos que muitos genes, talvez a maioria, são capazes de sofrer mutação para um estado letal. Assim, uma grande fração de todos os genes são absolutamente essenciais para a vida.

Letais dominantes que atuam cedo na vida são perdidos uma geração após ocorrerem porque os indivíduos que as levam morrem; entretanto, letais dominantes que atuam mais tarde na vida, após a reprodução, podem ser transmitidos para a geração seguinte. Letais recessivos podem permanecer um longo tempo em uma população porque podem ser escondidos na condição heterozigota por um alelo tipo selvagem. Mutações letais recessivas são detectadas observando-se proporções incomuns de segregação na prole de portadores heterozigotos. Um exemplo é a mutação *amarelo-letal*, A^Y , em camundongo (Fig. 4.7). Esta mutação é um dominante visível, que faz com que a pelagem seja amarela em vez de cinza-marrom (a cor tipo selvagem, também conhecida como *agouti*, que é determinada pelo alelo A^+). Além disso, entretanto, a mutação A^Y é um letal recessivo, matando os homozigotos $A^Y A^Y$ no início de seu desenvolvi-

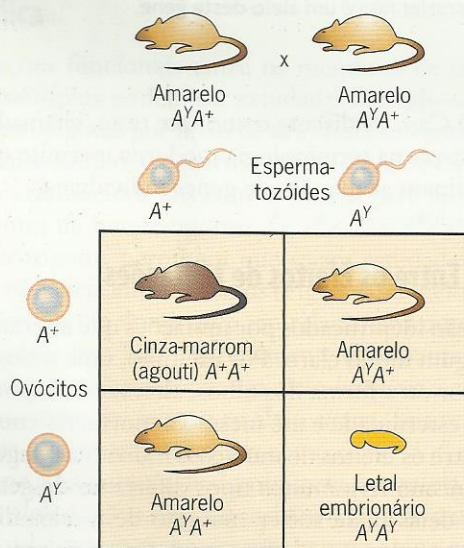


Fig. 4.7 ■ A^Y , a mutação *amarelo letal* em camundongos: um dominante visível que é também um recessivo letal. Um cruzamento entre portadores desta mutação produz heterozigotos amarelos e homozigotos cinza-marrom (agouti) em uma proporção de 2:1. Os homozigotos amarelos morrem ainda como embriões.

mento. Um cruzamento entre heterozigotos $A^Y A^+$ produz dois tipos de prole viável, amarelos ($A^Y A^+$) e cinza-marrom ($A^+ A^+$), em uma proporção de 2:1. Em uma prole suficientemente grande, esta proporção pode ser facilmente distinguida da proporção 3:1 que seria obtida se A^Y fosse apenas uma mutação dominante visível.

Os geneticistas usam convenções diferentes para representar genes e suas mutações. Mendel começou a praticar usando letras para indicar genes. Entretanto, ele começou simplesmente com a letra A e continuou com o alfabeto como símbolos quando necessário para representar genes em seus cruzamentos. William Bateson foi a primeira pessoa a usar letras mnemonicamente para simbolizar genes. Para o símbolo, Bateson escolheu a primeira letra da palavra que descrevia o efeito fenotípico do gene – assim, B , para um gene que causava flores azuis (*blue*), L para um gene que causava grãos de pólen longos. À medida que o número de genes conhecidos cresceu, tornou-se necessário usar duas ou mais letras para representar novos genes descobertos. As variações nos procedimentos para denominar e representar genes são discutidas no Enfoque Técnico: Símbolos Genéticos.

Os Genes Funcionam Produzindo Polipeptídeos

A extensa variação revelada por mutações indica que os organismos contêm muitos genes diferentes, e esses genes podem existir em vários estados alélicos. Entretanto, não nos diz como os genes de fato atuam no fenótipo. Como um gene influencia uma característica tal como cor dos olhos, textura da semente ou altura da planta?

Os primeiros geneticistas tiveram de responder a esta pergunta. Entretanto, hoje é claro que a maioria dos genes especifica um produto que afeta subsequentemente o fenótipo. Esta idéia, que foi discutida no livro de Bateson e foi apoiada pelas pesquisas de muitos cientistas, incluindo, e notadamente, o médico britânico *Sir Archibald Garrod* (veja Um Marco na

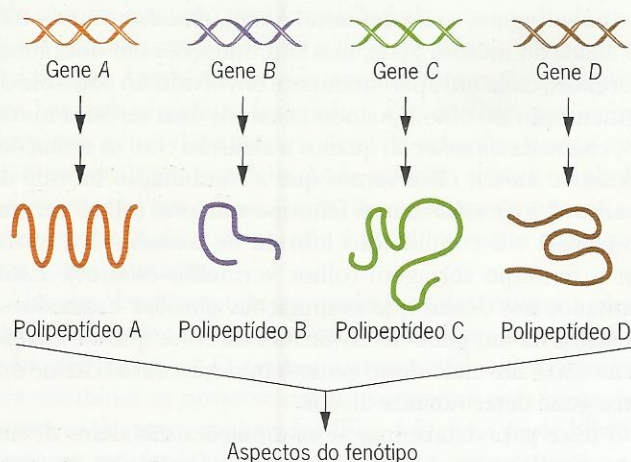


Fig. 4.8 ■ Relações entre genes e polipeptídeos. Cada gene especifica um polipeptídeo diferente. Esses polipeptídeos então agem influenciando o fenótipo do organismo.



ENFOQUE TÉCNICO: SÍMBOLOS GENÉTICOS

William Bateson começou a praticar a escolha dos símbolos genéticos mnemonicamente. Ao discutir o trabalho de Mendel, por exemplo, ele representou o alelo dominante para plantas altas de ervilha como *T* e o alelo recessivo para plantas baixas como *t*. Depois, tornou-se costume escolher os símbolos alélicos com base na característica mutante, e tais símbolos foram mudados para *D* (altas) e *d* (baixas). Esta convenção forneceu uma notação simples e consistente na qual os alelos dominante e recessivo de um determinado gene eram representados por uma única letra, e esta letra era mnemônica para a característica influenciada pelo gene. Bateson também criou os termos *genética*, *alelomorfo* (que depois foi reduzido para *alelo*), *homozigoto* e *heterozigoto* e introduziu a prática de representar as gerações em um esquema de reprodução como *P*, *F*₁, *F*₂, e assim em diante.

O sistema de denominação de genes que Bateson desenvolveu funcionou bem até que o número de genes que foram identificados excedeu a capacidade do alfabeto inglês. A partir de então, tornou-se necessário usar duas ou mais letras para simbolizar um gene. Por exemplo, um determinado alelo mutante em *Drosophila* faz com que os olhos sejam carmim em vez de vermelhos. Quando este alelo foi descoberto, recebeu o símbolo *cm* pois a única letra *c* já tinha sido usada para representar um alelo mutante que faz com que as asas sejam curvas em vez de retas. Hoje em dia, com milhares de genes identificados, geralmente é necessário usar três ou quatro letras, ou combinações de letras e números, para representar os genes. Por exemplo, mutações no gene *cnp* de *Drosophila* fazem com que as asas sejam enrugadas (*crumpled*) e mutações nos genes *Sh1* e *Sh2* de milho fazem com que as espigas sejam murchas (*shrunken*).

A descoberta de alelos múltiplos tornou a notação genética ainda mais complicada. Como as letras maiúsculas e minúsculas não eram mais adequadas para distinguir os alelos, os geneticistas começaram a combinar um símbolo gênico básico com um símbolo de identificação. Os geneticistas de *Drosophila* foram os primeiros a aplicar este procedimento. Eles tornaram o símbolo de identificação um expoente do símbolo do gene básico. Geralmente, tanto o símbolo do gene quanto o expoente têm algum significado mnemônico. Assim, por exemplo, *cn*² foi usado como símbolo do segundo alelo da cor de olho cinnabar que foi descoberto em *Drosophila*; e *ey*^D foi usado para representar um

alelo dominante que faz com que as drosófilas não tenham olhos (*eyeless*). Esta convenção foi ampliada para outros animais experimentais, tais como coelhos e camundongos. Por exemplo, *c*^{ch} foi usado para representar o alelo chinchilla do gene que determina se a pelagem do coelho é colorida ou incolor. Geneticistas de plantas adotaram uma variação desta prática. Eles usaram símbolos com hífen para identificar alelos mutantes. Por exemplo, *sh2-6801* representa um alelo mutante do gene *Sh2* que foi descoberto em 1968.

À medida que a nomenclatura genética se desenvolveu, tornou-se necessário usar um símbolo especial para representar o alelo tipo selvagem. Os primeiros geneticistas de *Drosophila* propuseram usar um sinal (+), às vezes escrito como expoente do símbolo do gene básico (por exemplo, *c*⁺). Esta notação simples passa a idéia de que o alelo tipo selvagem é o alelo padrão, ou normal, do gene, e é amplamente usado hoje em dia. Entretanto, persistem outras práticas para denominar os genes. Os geneticistas de plantas tendem a usar o próprio símbolo do gene para representar o alelo tipo selvagem, mas com destaque, e a primeira letra é maiúscula. Assim, *Sh2* é o alelo tipo selvagem do segundo gene para murcha que foi descoberto em milho, enquanto *sh2* é um alelo mutante.

A nomenclatura genética foi ainda mais complicada pela descoberta de genes por meio dos polipeptídeos que eles especificam. Tais descobertas introduziram símbolos de genes que são mnemônicos para produtos gênicos polipeptídicos. Por exemplo, o gene humano que especifica o polipeptídeo hipoxantina-guanina-fosforibosiltransferase é representado por *HPRT* e o gene de plantas que especifica o polipeptídeo álcool-desidrogenase é representado por *Adh*. Se letras maiúsculas são usadas nos símbolos gênicos ou apenas para a primeira letra, isto depende do organismo.

Hoje em dia existem muitos sistemas especializados para representar genes e alelos. Pesquisadores que trabalham com organismos diferentes – *Drosophila*, camundongos, plantas ou humanos – falam linguagens um pouco diferentes. Mais adiante, veremos ainda outros dialetos genéticos que foram criados para descrever os genes de vírus, bactérias e fungos. Estes sistemas diferentes de nomenclatura indicam que os símbolos em genética se desenvolveram em resposta a novas descobertas, uma evidência visível do crescimento em uma ciência jovem e dinâmica.

Genética: Garrod e os Erros Hereditários do Metabolismo, mais adiante neste capítulo), surgiu na metade do século vinte quando George Beadle e Edward Tatum descobriram que os produtos dos genes são *polipeptídeos* (Fig. 4.8).

Polipeptídeos são macromoléculas constituídas de uma cadeia linear de *aminoácidos*. Cada organismo faz milhares de polipeptídeos diferentes, cada um caracterizado por uma sequência específica de aminoácidos. Esses polipeptídeos são os constituintes fundamentais de *proteínas*. Dois ou mais po-

lipeptídeos podem combinar-se para formar uma proteína. Algumas proteínas, chamadas *enzimas*, funcionam como catalisadores em reações bioquímicas; outras formam componentes estruturais das células; e ainda outras são responsáveis por transportar substâncias dentro das células e entre elas. Beadle e Tatum propuseram que cada gene é responsável pela síntese de um determinado polipeptídeo. Quando um gene sofre mutação, seu produto polipeptídico ou não é feito ou está alterado, de tal modo que seu papel no organismo é

modificado. As mutações que eliminam ou alteram um polipeptídeo em geral estão associadas a um efeito fenotípico. Se este efeito é dominante ou recessivo irá depender da natureza da mutação. No Cap. 12, consideramos os detalhes sobre como os genes produzem polipeptídeos e no Cap. 13 discutiremos a base molecular da mutação.

Por que Algumas Mutações São Dominantes e Outras Recessivas?

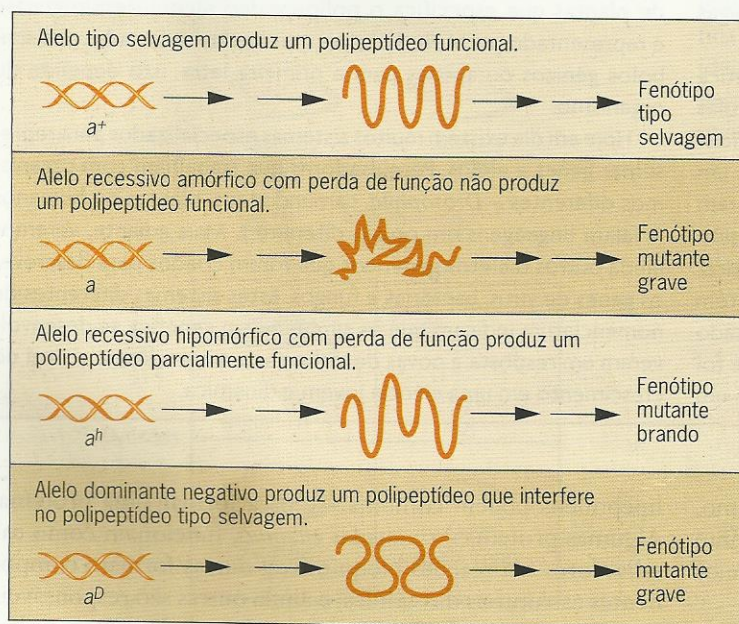
A descoberta de que os genes especificam polipeptídeos permitiu a compreensão da natureza das mutações dominantes e recessivas. Mutações dominantes têm efeitos fenotípicos em heterozigotos bem como em homozigotos, enquanto mutações recessivas têm estes efeitos apenas em homozigotos. O que responde por esta marcante diferença na expressão?

Mutações recessivas em geral envolvem perda de função do gene, isto é, quando o gene não especifica mais um polipeptídeo ou quando ele especifica um polipeptídeo não funcional ou subfuncional (Fig. 4.9). As mutações recessivas são portanto tipicamente alelos de **perda de função**. Tais alelos têm pouco ou nenhum efeito discernível na condição heterozigota com um alelo tipo selvagem porque o alelo selvagem especifica um polipeptídeo funcional que fará seu papel normal no organismo. O fenótipo de um heterozigoto mutante/selvagem será portanto o mesmo, ou essencialmente o mesmo, que o do homozigoto tipo selvagem. A mutação *cinnabar* em *Drosophila* é um exemplo de alelo de perda de função recessivo. O alelo tipo selvagem do gene

cinnabar produz um polipeptídeo que funciona como uma enzima na síntese do pigmento marrom que é depositado nos olhos de *Drosophila*. As moscas que são homozigotas para uma mutação de perda de função no gene *cinnabar* não podem produzir esta enzima e conseqüentemente não produzem nenhum pigmento marrom em seus olhos. O fenótipo dos mutantes homozigotos *cinnabar* é vermelho-claro, a cor do mineral cinábrio, donde o nome do gene. Entretanto, as moscas que são heterozigotas para a mutação *cinnabar* e seu alelo tipo selvagem têm olhos vermelhos, isto é, elas são fenotipicamente idênticas ao tipo selvagem. Nestas moscas, o alelo de perda de função é recessivo em relação ao alelo tipo selvagem porque este último produz enzima suficiente para sintetizar quantidades normais do pigmento marrom. A mutação *scarlet* já mencionada neste capítulo também é um exemplo de alelo recessivo de perda de função.

Algumas mutações recessivas resultam em perda parcial da função do gene. Por exemplo, o alelo *himalaio* do gene da cor de pelagem em mamíferos tais como coelhos e gatos especifica um polipeptídeo que funciona apenas nas partes do corpo onde a temperatura é reduzida. Esta perda parcial de função explica por que animais homozigotos para o alelo *himalaio* têm pêlos pigmentados em suas extremidades – cauda, pernas, orelhas e ponta do nariz – mas não no resto do corpo. Nas extremidades, o polipeptídeo especificado por este alelo é funcional, enquanto no resto do corpo ele não é. A expressão do alelo *himalaio* é, portanto, sensível a temperatura.

Algumas mutações dominantes também podem envolver perda de função do gene. Se o fenótipo controlado por um gene é sensível à quantidade do produto gênico, uma muta-



(a)

Genótipo	Poli-peptídeos presentes	Fenótipo	Natureza do alelo mutante
$a^+ a$		Tipo selvagem	Recessivo
$a^+ a^h$		Tipo selvagem	Recessivo
$a^+ a^D$		Mutante	Dominante

(b)

Fig. 4.9 ■ Diferenças entre mutações recessivas de perda de função e mutações dominantes de ganho de função. (a) Produtos polipeptídicos de mutações recessivas e dominantes. (b) Fenótipos de heterozigotos portando um alelo tipo selvagem e tipos diferentes de alelos mutantes.

ção de perda de função pode levar a um fenótipo mutante na condição heterozigota com um alelo tipo selvagem. Em tais casos, o alelo tipo selvagem, em si, não é capaz de gerar produto gênico suficiente para uma função completa normal. Com efeito, a mutação de perda de função reduz o nível do produto gênico abaixo do nível que é necessário para o fenótipo selvagem.

Outras mutações dominantes interferem na função do alelo tipo selvagem especificando polipeptídeos que inibem, antagonizam ou limitam a atividade do polipeptídeo tipo selvagem (Fig. 4.9). Tais mutações às vezes são chamadas de mutações **dominantes negativas**. Algumas das mutações do gene *T* no camundongo são exemplos de mutações dominantes negativas. Nós já vimos que na condição heterozigota essas mutações causam um encurtamento da cauda. Na condição homozigota, elas são letais porque o tronco do animal falha em se desenvolver totalmente. O gene *T* é portanto crítico para a formação de uma parte substancial do corpo do camundongo. A primeira mutação *T* dominante foi descoberta em 1927 por N. Dobrovolskaia-Zavadskaia, que trabalhou no Institute Pasteur em Paris. Muitas outras mutações *T* dominantes foram identificadas desde então, e recentemente várias foram analisadas em nível molecular. O produto polipeptídico do alelo *T* tipo selvagem liga-se a regiões específicas do DNA onde regula a expressão de genes que estão envolvidos no desenvolvimento do camundongo. O produto polipeptídico de um alelo *T* dominante negativo é um pouco mais curto que o polipeptídeo do tipo selvagem, e em heterozigotos ele interfere na função do polipeptídeo tipo selvagem. Esta interferência é suficientemente ampla para evocar um fenótipo totalmente sem cauda em camundongos que são heterozigotos para uma mutação *T* dominante negativa.

Algumas mutações dominantes causam um fenótipo mutante na condição heterozigota com um alelo tipo selvagem porque elas acentuam a função do produto gênico. A função acentuada pode surgir porque a mutação especifica um polipeptídeo novo ou porque faz com que o polipeptídeo tipo selvagem seja produzido onde ou quando ele não deveria ser. As mutações dominantes que funcionam deste modo são chamadas mutações de **ganho de função**. Em *Drosophila*, a mutação conhecida como *Antennapedia* (*Antp*) é uma mutação dominante de ganho de função. Na condição heterozigota com um alelo tipo selvagem, *Antp* faz com que pernas desenvolvam-se no lugar de antenas na cabeça da mosca. O motivo para esta bizarra transformação anatômica é que a mutação *Antp* faz com que o produto polipeptídico do gene *Antennapedia* seja produzido na cabeça, onde, normalmente, não é produzido. O produto do gene *Antennapedia* expandiu, assim, o domínio de sua função.

■ PONTOS IMPORTANTES

- Os genes em geral têm vários alelos.
- Alelos mutantes podem ser dominantes, recessivos, incompletamente dominantes ou co-dominantes.
- Se um híbrido que herda uma mutação recessiva de cada um de seus genitores tem um fenótipo mutante, então as muta-

ções recessivas são alelos do mesmo gene; se o híbrido tem um fenótipo selvagem, então as mutações recessivas são alelos de genes diferentes.

- A maioria dos genes codificam polipeptídeos.
- Na condição homozigota, as mutações recessivas em geral abolem ou diminuem a atividade do polipeptídeo.
- Algumas mutações dominantes produzem um polipeptídeo que interfere na atividade do polipeptídeo codificado pelo alelo tipo selvagem de um gene.

AÇÃO GÊNICA: DO GENÓTIPO AO FENÓTIPO

Os fenótipos dependem de fatores tanto genéticos quanto ambientais.

No começo do século vinte, os geneticistas tinham idéias imprecisas sobre como os genes produzem determinados fenótipos. Eles não sabiam nada sobre a química da estrutura ou função do gene nem tinham desenvolvido técnicas para estudá-las. Tudo que eles propuseram sobre a natureza da ação gênica foi deduzido da análise de fenótipos. Tais análises mostraram que os genes não atuam isoladamente. Em vez disso, eles atuam no contexto de um ambiente e às vezes atuam em conjunto com outros genes. Estas análises também mostraram que um determinado gene pode influenciar muitas características diferentes.

Influência do Ambiente

Um gene deve funcionar no contexto de ambientes tanto biológicos quanto físicos. Os fatores no ambiente físico são mais fáceis de estudar, pois determinados genótipos podem ser criados no laboratório sob condições controladas, permitindo uma avaliação dos efeitos de temperatura, luz, nutrição e umidade. Como um exemplo, consideremos a mutação de *Drosophila* conhecida como *shibire*. Na temperatura normal de cultura, 25°C, as moscas *shibire* são viáveis e férteis, mas são extremamente sensíveis a um choque súbito. Quando uma cultura *shibire* é agitada, as moscas, temporariamente paralisadas, caem para o fundo da garrafa de cultura. *Shibire* é a palavra para "paralisia" em japonês. Entretanto, se uma cultura de moscas *shibire* é colocada em temperatura um pouco mais alta, 29°C, todas as moscas caem para o fundo e morrem, mesmo sem um choque. Portanto, o fenótipo das mutações *shibire* é sensível a temperatura. A 25°C, a mutação é viável, mas a 29°C ela é letal. Uma explicação plausível é que a 25°C o gene mutante faz uma proteína parcialmente funcional, mas a 29°C esta proteína é totalmente não funcional.

Efeitos Ambientais na Expressão de Genes Humanos

As pesquisas de genética humana fornecem um exemplo de como o ambiente físico influencia um fenótipo. **Fenilceto-**

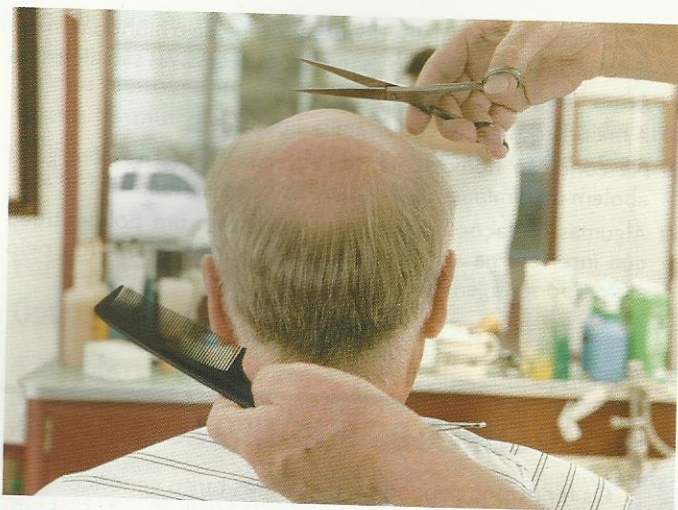


Fig. 4.10 ■ Homem com calvície prematura.

núria (PKU) é um distúrbio recessivo do metabolismo de aminoácidos. Crianças homozigotas para o alelo mutante acumulam substâncias tóxicas em seus cérebros. Embora não letais, essas substâncias podem prejudicar a capacidade mental afetando o desenvolvimento do cérebro. Os aspectos prejudiciais da PKU são relacionáveis a um determinado aminoácido, a fenilalanina, que é ingerida na dieta. Embora em si não tóxica, a fenilalanina é metabolizada em outras substâncias que são. Crianças com PKU que são alimentadas por dietas normais ingerem fenilalanina suficiente para causar as piores manifestações da doença. Entretanto, crianças que são alimentadas com dietas pobres em fenilalanina geralmente amadurecem sem os graves prejuízos mentais. Como a PKU pode ser diagnosticada em neonatos, o impacto clínico desta doença pode ser reduzido se as crianças que são homozigotas para PKU forem colocadas em uma dieta pobre em fenilalanina logo após o nascimento. Este exemplo ilustra como um fator ambiental, a dieta, pode ser manipulado para modificar um fenótipo que de outro modo se tornaria uma tragédia.

O ambiente biológico também pode influenciar a expressão fenotípica de genes. O **padrão de calvície** em humanos é um exemplo bem conhecido (Fig. 4.10). Aqui, o padrão

biológico relevante é o sexo. O padrão de calvície prematura se deve a um alelo que é expresso diferencialmente nos dois sexos. Em homens, tanto os homozigotos quanto os heterozigotos para este alelo desenvolvem calvície, enquanto em mulheres, apenas as homozigotas tendem a se tornar calvas, sendo isto geralmente limitado a afinamento do cabelo. A expressão deste alelo é provavelmente disparada pelo hormônio masculino **testosterona**. As mulheres produzem muito menos deste hormônio e estão, portanto, em risco muito raramente afetado de desenvolver áreas de calvície. A natureza influenciada pelo sexo do padrão de calvície mostra que fatores biológicos podem controlar a expressão dos genes.

Penetrância e Expressividade

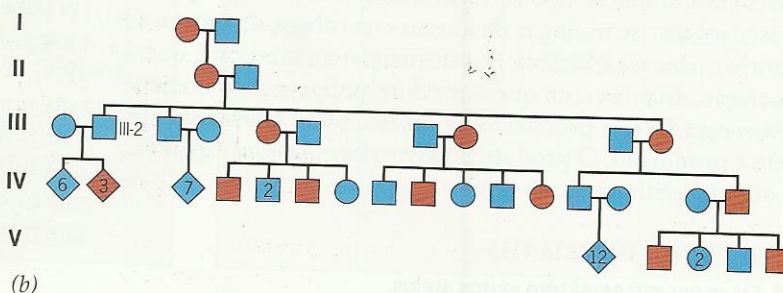
Quando as pessoas não apresentam uma característica muito embora tenham o genótipo apropriado, diz-se que a característica exibe **penetrância incompleta**. Um exemplo de penetrância incompleta em humanos é a **polidactilia**, a presença de dedos e artelhos extras (Fig. 4.11a). Esta condição se deve a uma mutação dominante, *P*, que só se manifesta em alguns de seus portadores. No heredograma da Fig. 4.11b, a pessoa III-2 deve ser um portador muito embora não tenha dedos ou artelhos extras. O motivo é que tanto sua mãe quanto três de seus filhos são polidáctilos, uma indicação da transmissão da mutação por III-2. A penetrância incompleta pode ser um grave problema na análise de um heredograma porque pode resultar na atribuição incorreta dos genótipos.

O termo **expressividade** é usado se uma característica não se manifesta uniformemente entre os afetados que a apresentam. A mutação dominante ocular *Lobe* (Fig. 4.12) em *Drosophila* é um exemplo. O fenótipo associado a esta mutação é extremamente variável. Algumas moscas heterozigotas têm pequenos olhos compostos, enquanto outras têm olhos grandes lobulados. Entre estes extremos, há toda uma gama de fenótipos. A mutação *Lobe* é portanto dita de **expressividade variável**.

A penetrância incompleta e a expressividade variável indicam que a via entre um genótipo e seus fenótipos é objeto de considerável modulação. Os geneticistas sabem que parte dessa modulação é decorrente de fatores ambientais,



(a)



(b)

Fig. 4.11 ■ Polidactilia em seres humanos. (a) Fenótipo mostrando dedos extras. (b) Heredograma mostrando a herança desta característica dominante incompletamente penetrante.

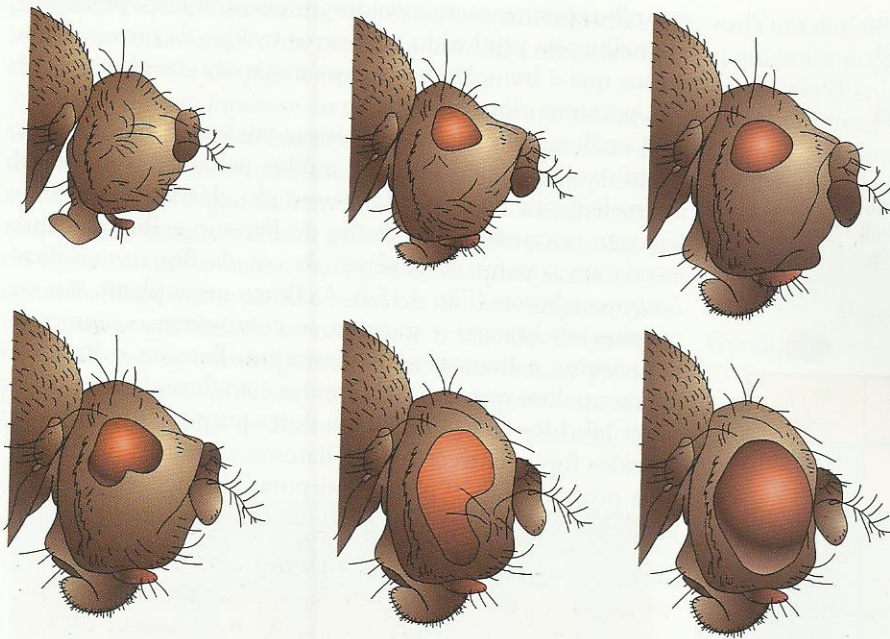


Fig. 4.12 ■ Expressividade variável da mutação *Lobe* em *Drosophila*. Cada mosca é heterozigota para esta mutação dominante; entretanto, os fenótipos variam de ausência completa de olho a um olho quase tipo selvagem.

mas parte também se deve a fatores no ambiente genético. A evidência clara de tais fatores vem de experimentos de cruzamento mostrando que dois ou mais genes podem afetar determinada característica.

Interações Gênicas

Algumas das primeiras evidências de que uma característica pode ser influenciada por mais de um gene foram obtidas por Bateson e Punnett de experimentos de cruzamento em galinhas. Seu trabalho foi feito logo após o redescobrimto das publicações de Mendel. Tipos diferentes de galinhas domésticas têm formas diferentes de cristas (Fig. 4.13). As Wyandottes têm cristas “*rose*”, as Brahmas têm cristas “*pea*” e as Leghorns têm cristas “*single*”. Cruzamentos entre Wyandottes e Brahmas produzem galinhas que têm um outro tipo de crista, chamada “*walnut*”. Bateson e Punnett descobriram que o tipo de crista é determinado por dois genes que segregam independentemente, *R* e *P*, cada um dos quais com dois alelos (Fig. 4.14). Wyandottes (com cristas *rose*) têm o genótipo *RR pp* e Brahmas (com cristas *pea*) têm o genótipo *rr PP*. Os híbridos de F_1 entre estas duas variedades são portanto *Rr Pp*, e fenotipicamente eles têm cristas *walnut*. Se tais híbridos forem entrecruzados, todos os quatro tipos de cristas aparecerão na prole: 9/16 *walnut* (*R- P-*), 3/16 *rose* (*R- pp*), 3/16 *pea* (*rr P-*) e 1/16 *single* (*rr pp*). A linhagem Leghorn, que tem crista tipo *single*, deve, portanto, ser homozigota para ambos os alelos recessivos.

O trabalho de Bateson e Punnett demonstrou que dois genes que segregam independentemente podem afetar uma característica. Combinações diferentes de alelos dos dois genes resultaram em fenótipos diferentes, supostamente devido a interações de seus produtos em nível bioquímico ou celular.

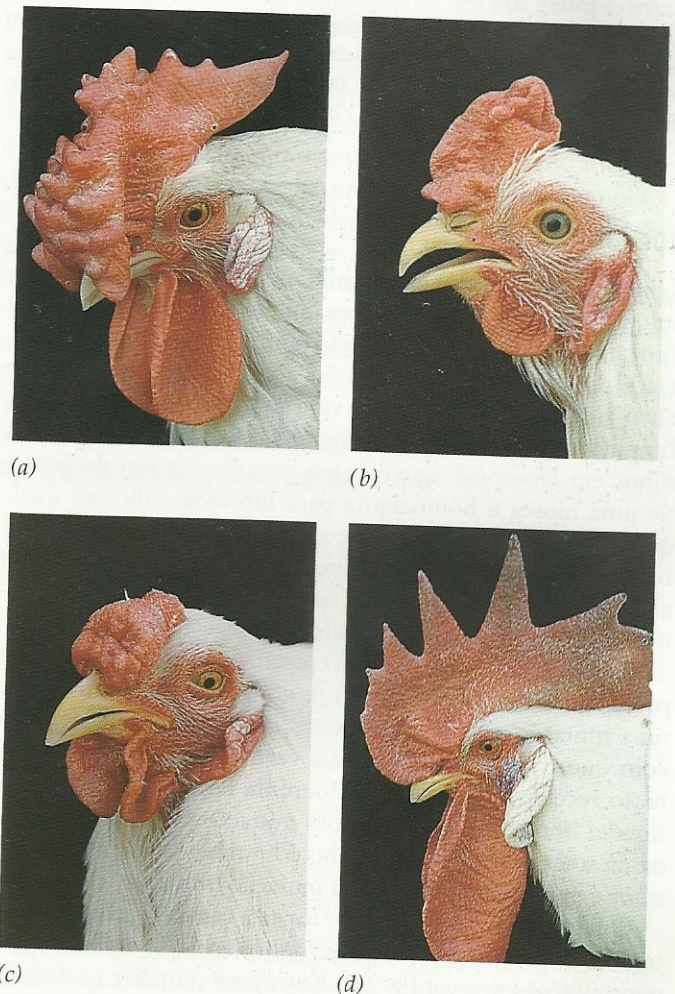


Fig. 4.13 ■ Formas das cristas de galinha de raças diferentes. (a) *Rose*, Wyandottes; (b) *pea*, Brahmas; (c) *walnut*, híbrida do cruzamento entre galinhas com cristas *rose* e *pea*; (d) *single*, Leghorns.

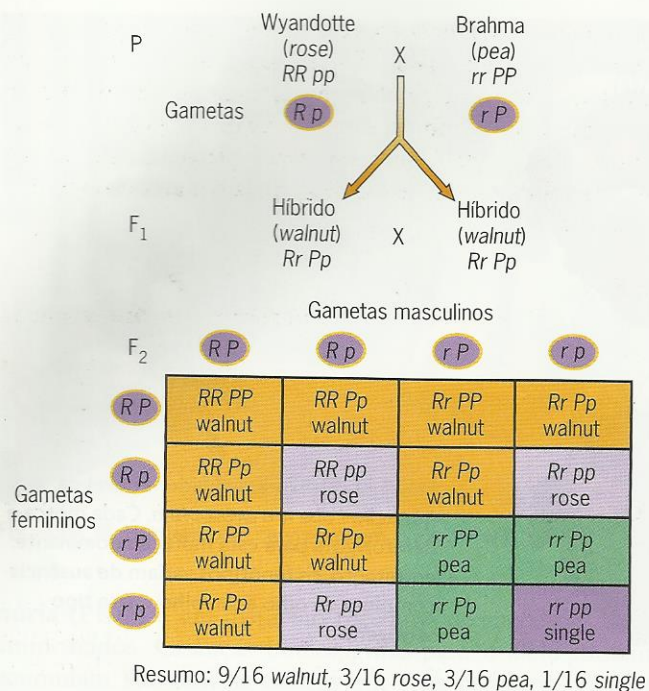


Fig. 4.14 ■ O experimento de Bateson e Punnett sobre a forma das cristas em galinhas. O entrecruzamento na F_1 produz quatro fenótipos, cada um destacado por uma cor diferente no quadrado de Punnett, em uma proporção 9:3:3:1.

Epistasia

Quando dois ou mais genes influenciam uma característica, um alelo de uma delas pode ter um efeito predominante sobre o fenótipo. Quando um alelo tem tal efeito sobrepujante, ele é dito epistático em relação a outros genes que estão envolvidos; o termo **epistasia** vem da palavra grega para “sobrepujante”. Por exemplo, sabemos que a pigmentação dos olhos em *Drosophila* envolve um grande número de genes. Se uma mosca é homozigota para um alelo nulo em qualquer um destes genes, a via de síntese do pigmento pode ser bloqueada, e será produzida uma cor de olho anormal. Este alelo essencialmente anula o trabalho de todos os outros genes, mascarando suas contribuições para o fenótipo.

Um alelo mutante é dito epistático quando encobre a presença de outro alelo mutante no genótipo. Já vimos que uma mutação recessiva no gene *cinnabar* de *Drosophila* faz com que o olho da mosca seja vermelho-claro. Uma mutação recessiva em um gene diferente faz com que os olhos fiquem brancos. Quando ambas mutações são homozigotas na mesma mosca, a cor do olho é branca. Assim, a mutação *white* é epistática em relação à mutação *cinnabar*.

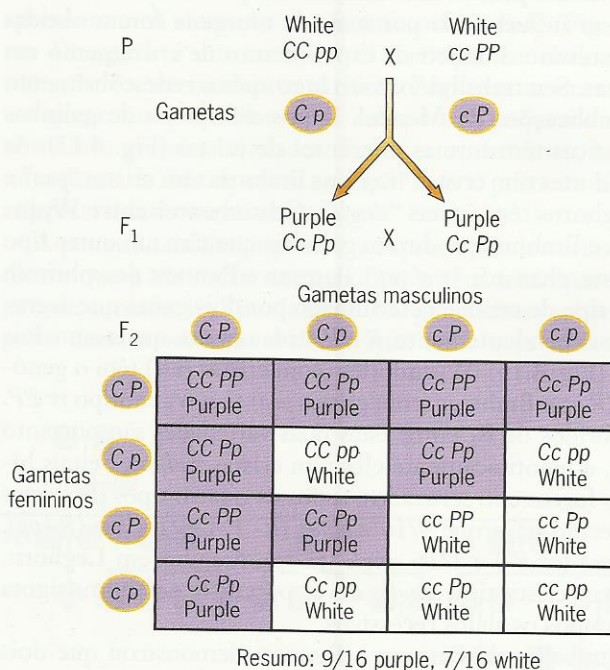
Que mecanismo fisiológico torna a mutação *white* epistática em relação à *cinnabar*? Por muitos anos, a resposta a esta pergunta não era conhecida. Entretanto, análise molecular recente mostrou que o produto polipeptídico do gene *white* transporta pigmento para o olho da *Drosophila*. Quando este gene sofre mutação, o polipeptídeo transportador não é feito,

e os olhos permanecem incolores, muito embora o pigmento vermelho seja produzido em outros tecidos da mosca. Uma mosca que é homozigota para as mutações *cinnabar* e *white* tem portanto olhos brancos.

A análise das relações epistáticas tais como aquela entre *cinnabar* e *white* pode sugerir modos pelos quais os genes controlam um fenótipo. Um exemplo clássico desta análise vem novamente do trabalho de Bateson e Punnett, que estudaram o controle genético da cor da flor nas ervilhas *Lathyrus odoratus* (Fig. 4.15a). As flores nesta planta são ou púrpura ou brancas – púrpura se contiverem o pigmento antocianina e brancas se não tiverem. Bateson e Punnett cruzaram duas variedades diferentes com flores brancas para obter híbridos de F_1 , todas com flores púrpura. Quando tais híbridos foram entrecruzados, Bateson e Punnett obtiveram uma proporção de 9 plantas púrpura:7 brancas na F_2 . Eles explicaram os resultados propondo que dois genes se-



(a)



(b)

Fig. 4.15 ■ (a) Flores brancas e púrpura de ervilhas. (b) O experimento de Bateson e Punnett sobre o controle genético de cor da flor em ervilhas.

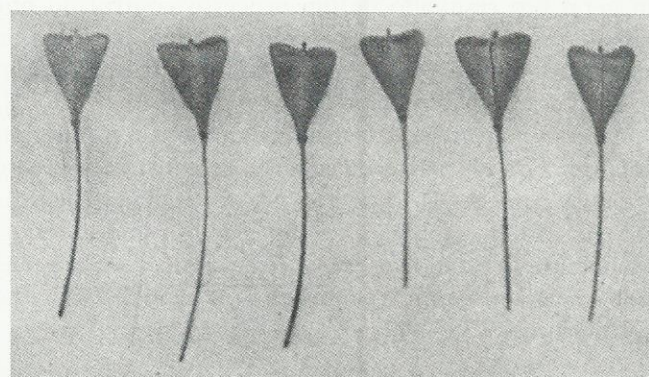
gregam independentemente, *C* e *P*, estão envolvidos na síntese de antocianina, e cada gene tem um alelo recessivo que abole a produção do pigmento (Fig. 4.15b).

Em vista desta hipótese, as variedades parentais teriam de ter genótipos complementares: *cc PP* e *CC pp*. Quando estas duas variedades foram cruzadas, elas produziram duplos heterozigotos *Cc Pp* que tinham flores púrpura. Neste sistema, um alelo dominante de cada gene é necessário para a síntese do pigmento antocianina. Na F_2 , 9/16 das plantas são *C-* *P-* e têm flores púrpura; os 7/16 restantes são homozigotos para pelo menos um dos alelos recessivos e têm flores brancas. Observe que os homozigotos duplos recessivos, *cc pp*, não são fenotipicamente diferentes dos homozigotos recessivos simples. O trabalho de Bateson e Punnett estabeleceu que

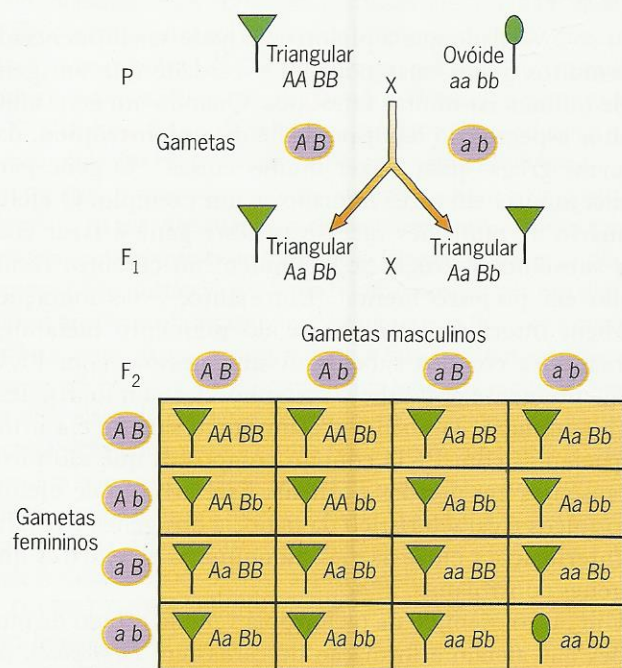
cada um dos alelos recessivos é epistático em relação ao alelo dominante do outro gene. Uma explicação plausível é que cada alelo dominante produz uma enzima que controla uma etapa na síntese de antocianina a partir de um precursor bioquímico. Se um alelo dominante não estiver presente, sua etapa na via de biossíntese é bloqueada e a antocianina não é produzida:

Gene	<i>C</i>	<i>P</i>
	Precursor → Intermediário	Intermediário → Antocianina
Genótipo		
<i>C-</i> <i>P-</i>	+	+
<i>cc P-</i>	+	-
<i>C- pp</i>	+	-

Outro estudo clássico de epistasia foi feito por George Shull usando uma planta chamada bolsa de pastor, *Bursa bursa-pastoris* (Fig. 4.16a). As cápsulas das sementes desta planta são ou triangulares ou de forma ovóide. As cápsulas ovóides são produzidas apenas se a planta for homozigota para os alelos recessivos dos dois genes, isto é, se ela tiver o genótipo *aa bb*. Se o alelo dominante de um dos genes estiver presente, a planta produz cápsulas triangulares. A evidência para esta conclusão vem de cruzamentos entre plantas duplamente heterozigotas (Fig. 4.16b). Tais cruzamentos produzem prole em uma proporção de 15 triangulares:1 ovóide, indicando que o alelo dominante de um gene é epistático em relação ao alelo recessivo do outro. Os dados sugerem que a forma da cápsula é determinada por vias do desenvolvimento duplicadas, qualquer uma delas podendo produzir uma cápsula triangular. Uma via envolve o alelo dominante do gene *A*, e a outra o alelo dominante do gene *B*. Uma substância precursora que pode ser convertida em um produto resulta em uma cápsula triangular de semente por uma destas vias. Apenas quando ambas as vias estão bloqueadas por alelos homozigotos recessivos é que o fenótipo triangular é suprimido e é produzida uma cápsula ovóide:



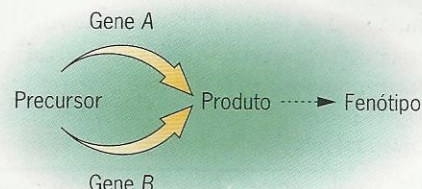
(a)



Resumo: 15/16 triangulares, 1/16 ovóide

(b)

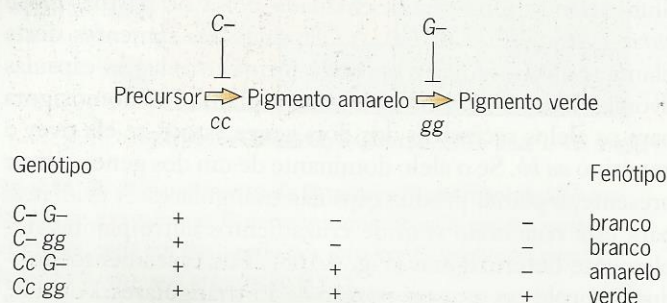
Fig. 4.16 ■ (a) Cápsulas triangulares de sementes de bolsa-de-pastor, *Bursa bursa-pastoris*. (b) Cruzamentos mostrando controle de genes duplicados de forma de cápsula de sementes na bolsa-de-pastor.



Genótipo	Gene A	Gene B	Fenótipo
<i>A- B-</i>	+	+	triangular
<i>aa B-</i>	+	+	triangular
<i>A- bb</i>	+	+	triangular
<i>aa bb</i>	+	-	ovóide

Em outros casos de epistasia, o produto de um gene pode inibir a expressão de outro gene. Considere, por exemplo, a herança da cor do fruto em abóbora. Plantas que possuem o alelo dominante *C* produzem fruto branco, enquanto plantas que são homozigotas para o alelo recessivo *c* produzem

fruto colorido. Se uma abóbora também for homozigota para o alelo recessivo *g* de um gene que segrega independentemente, o fruto será verde. Entretanto, se ela tiver o alelo dominante *G* deste gene, o fruto será amarelo. Estas observações sugerem que os dois genes controlam etapas na síntese do pigmento verde. A primeira etapa converte um precursor incolor em um pigmento amarelo, e a segunda etapa converte este pigmento amarelo em um pigmento verde. Se a primeira etapa é bloqueada (pela presença do alelo *C*), nenhum dos pigmentos será produzido e o fruto será branco. Se apenas a segunda etapa estiver bloqueada (pela presença do alelo *G*), o pigmento amarelo não poderá ser convertido em pigmento verde e o fruto será amarelo. Podemos resumir estas idéias com um diagrama que mostra o controle genético da síntese de pigmentos nesta via bioquímica:



As setas no diagrama mostram as etapas na via. O genótipo abaixo da seta indica a etapa que ocorre, enquanto o genótipo acima de uma seta inibe que esta etapa ocorra. É costume em genética representar o efeito inibidor de um genótipo desenhando uma barra em T partindo do genótipo para a etapa relevante na via. Neste exemplo, o alelo *C* inibe a primeira etapa e o alelo *G* inibe a segunda etapa. Devido a seu papel como um inibidor da primeira etapa, o alelo *C* é epistático em relação a ambos os alelos do outro gene. Independente de qual dos alelos deste outro gene esteja presente

em uma planta, o alelo *C* fará com que esta planta produza fruto branco.

A Fig. 4.17 mostra o resultado de um cruzamento entre plantas heterozigotas para ambos os genes determinantes da cor do fruto. Quando plantas *Cc Gg* são entrecruzadas, elas produzem uma prole que segrega três classes fenotípicas: branca, amarela e verde. A prole com o fruto verde é homozigota para os alelos recessivos de ambos os genes; isto é, elas são *cc gg*, e sua frequência é de 1/16. A prole com frutos amarelos é homozigota para *c* e possui pelo menos uma cópia de *G*. Sua frequência é de 3/16. A prole com frutos brancos possui pelo menos uma cópia de *C*, e o restante do genótipo não tem importância. A frequência das plantas com frutos brancos é de 12/16.

Estes exemplos indicam que um determinado fenótipo em geral é o resultado de um processo controlado por mais de um gene. Cada gene controla uma etapa na via que é parte do processo. Quando um gene sofre mutação para um estado não funcional ou parcialmente funcional, o processo pode ser perturbado, levando a um fenótipo mutante. Grande parte da análise genética moderna é devotada a investigação de vias envolvidas em processos biológicos importantes tais como metabolismo e desenvolvimento. O estudo das relações epistáticas entre genes pode ajudar a discernir o papel que cada gene desempenha nestes processos.

Pleiotropia

Não só é verdade que a pleiotropia pode ser influenciada por muitos genes, mas também é verdade que um gene pode influenciar muitos fenótipos. Quando um gene afeta muitos aspectos do fenótipo, ele é dito **pleiotrópico**, das palavras gregas para "fazer muitas coisas". O gene para fenilcetonúria em seres humanos é um exemplo. O efeito primário de mutações recessivas neste gene é fazer com que substâncias tóxicas se acumulem no cérebro, resultando em prejuízo mental. Entretanto, essas mutações também interferem na síntese do pigmento melanina, clareando a cor dos cabelos. Assim, pessoas com PKU freqüentemente têm cabelo castanho-claro ou louro. Testes bioquímicos também revelam que o sangue e a urina de pacientes com PKU contêm compostos que são raros ou ausentes em pessoas normais. Este arranjo de efeitos fenotípicos é característico da maioria dos genes e resulta de interligações entre as vias bioquímicas e celulares que os genes controlam.

Um outro exemplo de pleiotropia vem do estudo de mutações que afetam a formação de cerdas em *Drosophila*. As moscas tipo selvagem têm cerdas longas levemente curvas na cabeça e no tórax. As moscas homozigotas para a mutação de cerdas *singed* têm cerdas curtas torcidas nestas partes do corpo, como se tivessem sido chamuscadas. Assim, o produto do gene tipo selvagem *singed* é necessário para a formação adequada de cerdas. Ele também é necessário para

		White Cc Gg		X	White Cc Gg	
		Gametas masculinos				
		C G	C g	c G	c g	
Gametas femininas	C G	CC GG white	CC Gg white	Cc GG white	Cc Gg white	
	C g	CC Gg white	CC gg white	Cc Gg white	Cc gg white	
	c G	Cc GG white	Cc Gg white	cc GG yellow	cc Gg yellow	
	c g	Cc Gg white	Cc gg white	cc Gg yellow	cc gg green	

Fig. 4.17 ■ Segregação na prole de um cruzamento entre plantas de abóbora heterozigotas para dois genes que controlam a cor do fruto.



UM MARCO NA GENÉTICA: Garrod e os Erros Hereditários do Metabolismo

Quando as idéias de Mendel emergiram da obscuridade em 1900, elas rapidamente tiveram aplicação nas pesquisas do metabolismo humano. Desde o final do século dezenove, químicos e físicos estudaram como as moléculas ingeridas nos alimentos são processadas no corpo humano. A degradação dos alimentos é uma parte essencial do metabolismo. Ela fornece energia ao corpo para manter a vida e a matéria-prima para sintetizar novos compostos.

Grande parte dos primeiros trabalhos sobre metabolismo enfocaram produtos químicos de decomposição. As técnicas de química orgânica permitiram aos pesquisadores identificar esses materiais na urina e monitorar seus níveis após a ingestão de vários alimentos. Às vezes, a urina era coletada de pacientes em hospitais; às vezes era coletada dos próprios pesquisadores. Analisando os resultados destes experimentos de alimentação, os pesquisadores puderam relacionar dejetos químicos específicos a determinadas fontes alimentares. Também nesta época, experimentos com microrganismos, animais e plantas forneceram evidências de que os alimentos são decompostos de modo gradativo por uma série de reações, que hoje chamamos vias bioquímicas, e que cada reação é catalisada por uma determinada enzima.

O médico britânico Archibald E. Garrod (Fig. 1) juntou este tipo de análise bioquímica à emergente ciência da genética. Garrod ficou fascinado pelo que ele chamou de “erros inatos do metabolismo”, querendo dizer condições herdadas que “resultam da falha de alguma etapa ou outra na série de mudanças químicas que constituem o metabolismo”.¹ Garrod enfocou sua atenção em quatro destas condições: albinismo, alcaptonúria, cistinúria e pentosúria.

O albinismo surge de uma incapacidade de produzir o pigmento melanina. Pessoas com albinismo têm pele branca e cabelos claros, e seus olhos são rosa devido a ausência de qualquer pigmento na íris. Alcaptonúria, cistinúria e pentosúria são caracterizadas pela excreção de substâncias diferentes na urina. Na alcaptonúria, a substância é o ácido homogentísico, que nos dias de Garrod era chamada alcaptona. Esta substância fica preta quando exposta ao ar. Assim, a alcaptonúria é facilmente diagnosticada, mesmo em crianças, em que uma fralda molhada escurece, sendo um sinal claro da condição. Na cistinúria, o aminoácido cistina é excretado na urina. Caso a cistina fique muito concentrada, ela se cristaliza e forma cálculos, “pedras nos rins”, que são dolorosos e podem danificar os rins. Na pentosúria, o açúcar xilulose é excretado na urina. Ao contrário da cistinúria, nem a alcaptonúria nem a pentosúria estão associadas a uma patologia grave.

Em 1902, Garrod publicou um trabalho curto no periódico médico britânico *The Lancet*.² Esta publicação enfocou estudos de alcaptonúricos, que Garrod descobriu surgirem desproporcional-

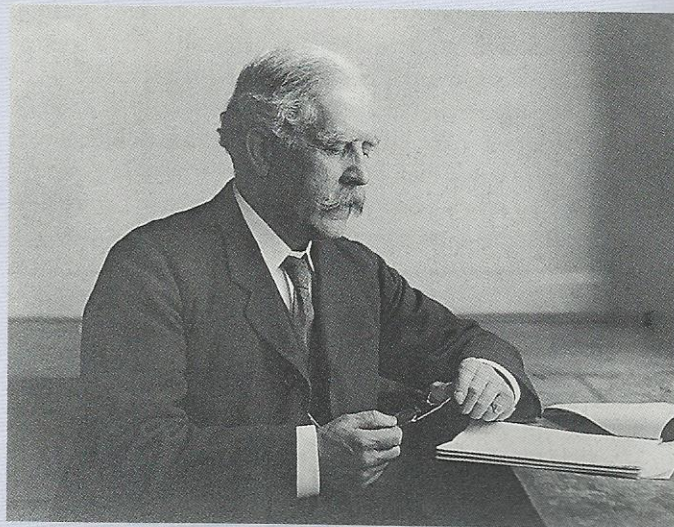


Fig. 1 ■ Archibald Garrod.

mente de genitores que eram primos de primeiro grau. As discussões com William Bateson, o grande campeão do mendelismo, deram a Garrod uma explicação plausível sobre esta observação. Garrod sugeriu que os alcaptonúricos eram homozigotos para um alelo mutante recessivo. Embora raro na população geral, esse alelo era mais provável estar em homozigose na prole de parentes do que na prole de pessoas não aparentadas. Em um livro baseado em palestras que ele apresentou no Royal College of Physicians em 1908, Garrod resumiu suas idéias deste modo:

Foi destacado por Bateson... que o modo de herança da alcaptonúria tem uma pronta explicação se a anomalia em questão for vista como uma característica recessiva rara no sentido mendeliano. A lei de Mendel diz que com relação a duas características mutuamente exclusivas, uma das quais tende a ser dominante e a outra recessiva, os indivíduos cruzados tendem a manifestar o caráter dominante, mas quando a prole dos híbridos é entrecruzada será exibida uma ou outra das características e consistirá em dominantes e recessivos em proporções definidas... Apenas quando dois gametas recessivos se juntam na fertilização, o indivíduo resultante apresenta a característica recessiva.

Se a característica recessiva for rara, muitas gerações podem ser puladas antes que ocorra a união de dois de tais gametas, pois as famílias nas quais são produzidos serão pouco numéricas, e a chance de que em um determinado casamento ambos os genitores contribuam com tais gametas será muito pequena. Quando, entretanto, ocorre o casamento entre dois membros de tal família, a chance será muito maior, e a prole de tal casamento provavelmente exibirá a peculiaridade.³

¹Garrod, A.E. 1909. *Inborn Errors of Metabolism*. Em *Garrod's Inborn Errors of Metabolism*, 1963, por H. Harris, Oxford University Press.

²Garrod, A.E. 1902. The incidence of alcaptonuria: a study of chemical individuality. *Lancet* ii, 1616-1620.

³Garrod, *Inborn Errors of Metabolism*.



UM MARCO NA GENÉTICA (continuação)

Assim, foi identificado o primeiro alelo recessivo em humanos. E mais, foi mostrado que este alelo está associado a uma anomalia metabólica. Garrod supôs que a alcaptonúria era causada por uma incapacidade de metabolizar o ácido homogentísico, que é então excretado na urina. Esta fusão da genética e da bioquímica levou à idéia de que o metabolismo está sob controle genético. Garrod também teve a evidência de que o albinismo ocorre desproporcionalmente entre a prole de primos de primeiro grau e é portanto causado por um alelo recessivo, mas, para a cistinúria e a pentosúria, ele só pôde especular sob uma base genética. Entretanto, com base em seu conhecimento detalhado da alcaptonúria, Garrod foi convencido de que os genes influenciam a constituição química das pessoas.

Hoje sabemos que o albinismo, a alcaptonúria e outro erro inato, a fenilcetonúria, são causados por bloqueios em vias bioquímicas que metabolizam dois aminoácidos, fenilalanina e tirosina, que são derivados de proteína ingerida nos alimentos (Fig. 2). Na fenilcetonúria, o aminoácido fenilalanina não pode ser convertido em tirosina devido a um defeito na enzima fenilalanina-hidroxilase. Em vez disso, a fenilalanina é convertida em outra substância, o ácido fenilpirúvico, que se acumula em níveis que irão danificar tecidos moles, especialmente os do cérebro. Se não for tratada com redução da ingestão de fenilalanina na dieta, esta anomalia metabólica pode prejudicar o desenvolvimento do cérebro. No albinismo, a via que converte tirosina no pigmento melanina é bloqueada por um defeito na enzima tirosinase. Consequentemente, o pigmento melanina não pode ser produzido. Na alcaptonúria, o bloqueio metabólico é em uma reação que converte ácido homogentísico em ácido maleilacetacético, o qual, por sua vez, é convertido por outras reações em ácido fumárico e ácido acetacético. Esse bloqueio, como Garrod corretamente indicou, faz com que o ácido homogentísico se acumule e seja excretado na urina. Pesquisas modernas mostraram que a pentosúria é causada por um bloqueio em uma via que degrada o açúcar xilulose. Entretanto, a cistinúria, a mais complexa das condições que Garrod estudou, não se deve a um bloqueio no metabolismo de cistina. Ela deriva da incapacidade dos túbulos renais de reabsorver cistina do líquido filtrado nos rins.

O trabalho de Garrod criou uma ligação crítica entre a genética e a bioquímica. Entretanto, levou mais de três décadas para que seu significado total fosse reconhecido. O evento esclarecedor surgiu em 1941, quando George Beadle e Edward Tatum demonstraram que os genes codificam as enzimas que catalisam reações bioquímicas. Com a descoberta de Beadle e Tatum, a era da genética bioquímica iniciou seu desenvolvimento.

QUESTÕES PARA DISCUSSÃO

1. O trabalho de Garrod exemplifica um enfoque interdisciplinar de pesquisa. Você consegue encontrar exemplos atuais de

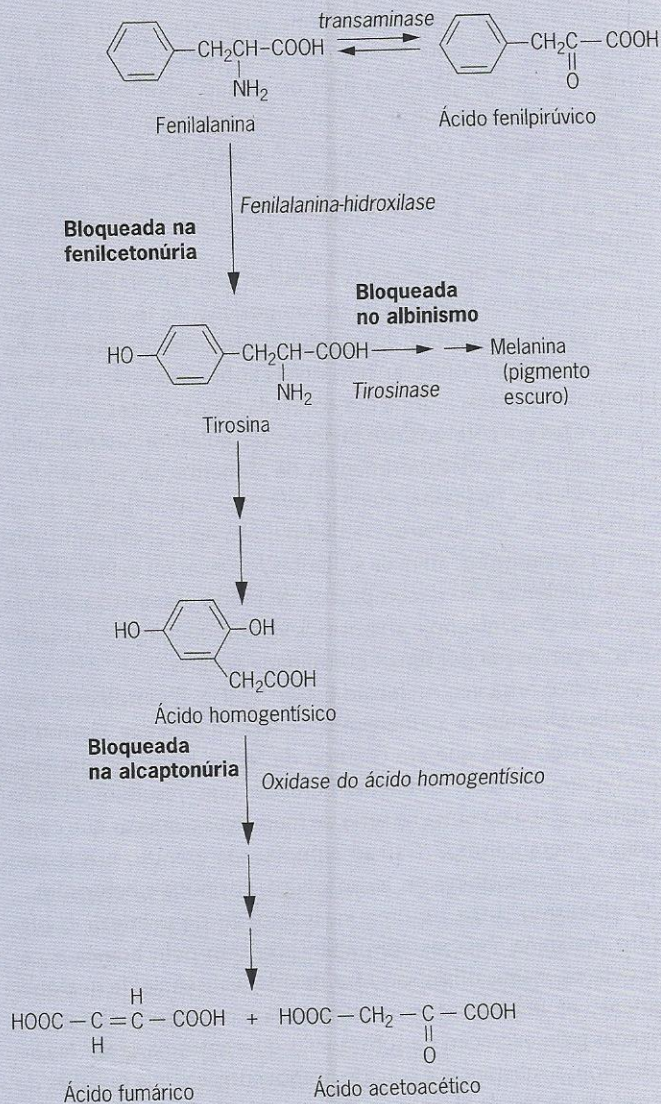


Fig. 2 ■ Vias bioquímicas envolvendo a modificação dos aminoácidos fenilalanina e tirosina.

projetos de pesquisa interdisciplinares? Existem dificuldades especiais associadas a estes tipos de projetos?

2. Muitos dos dados obtidos por Garrod e seus colegas vieram de experimentos com humanos. Substâncias químicas foram ministradas a pessoas para estudar aspectos do metabolismo. Outros tipos de dados vieram de estudos com animais tais como coelhos e cães. Que aspectos éticos estão envolvidos na pesquisa com humanos e animais? Como esta pesquisa é regulada hoje em dia?

a produção de ovócitos férteis saudáveis. Sabemos deste fato porque as fêmeas que são homozigotas para algumas mutações *singed* são totalmente estéreis. Elas põem ovos frágeis malformados que nunca eclodem. Entretanto, tais mutações não têm efeito adverso na fertilidade masculina. Assim, o gene *singed* controla pleiotropicamente a formação tanto de cerdas e ovos nas fêmeas quanto a formação de cerdas em machos.

PONTOS IMPORTANTES

- A ação gênica é afetada por fatores biológicos e físicos no ambiente.
- Dois ou mais genes podem influenciar uma característica.
- Um alelo mutante é epistático em relação a um alelo mutante de outro gene se tiver um efeito predominante no fenótipo.
- Um gene é pleiotrópico se ele influencia muitos fenótipos diferentes.

Exercícios Básicos Ilustram a análise genética básica.

1. Um pesquisador descobriu um novo sistema de grupo sanguíneo para seres humanos. O sistema envolve dois antígenos, P e Q, cada um determinado por um alelo diferente de um gene chamado *N*. Os alelos para estes antígenos são mais ou menos igualmente freqüentes na população geral. Se os alelos N^P e N^Q são co-dominantes, que antígenos devem ser detectados no sangue de heterozigotos $N^P N^Q$?

Resposta: Ambos os antígenos P e Q devem ser detectados porque a co-dominância significa que ambos os alelos em heterozigotos serão expressos.

2. A cor das flores em uma planta de jardim está sob o controle de um gene com vários alelos. Os fenótipos dos homozigotos e heterozigotos deste gene são os seguintes:

Homozigotos

WW	vermelha
ww	branca pura
$w^s w^s$	branca pontilhada de vermelho
$w^p w^p$	branca com manchas vermelhas regulares

Heterozigotos

W com qualquer outro alelo	vermelha
w^p com w^s ou w	branca com manchas vermelhas regulares
$w^s w$	branca pontilhada de vermelho

Coloque os alelos em uma hierarquia de dominância.

Resposta: W é dominante em relação a todos os outros alelos, w^p é dominante em relação a w^s e w e w^s são dominantes em relação a w . Assim, a hierarquia de dominância é $W > w^p > w^s > w$.

3. Duas linhagens de camundongo descobertas independentemente são homozigotas em relação a uma mutação recessiva que faz com que os olhos sejam pequenos; os fenótipos das duas linhagens são indistinguíveis. A mutação em uma linhagem é chamada *little eye* e a mutação na outra é chamada *tiny eye*. Uma terceira linhagem é heterozigota para uma mutação dominante que elimina os olhos; a mutação nesta linhagem é chamada *Eyeless*. Como você determinaria se as mutações *little eye*, *tiny eye* e *Eyeless* são alelos do mesmo gene?

Resposta: O procedimento para determinar se duas mutações recessivas são alelos do mesmo gene é cruzar seus res-

pectivos homozigotos para obter prole híbrida e então avaliar o fenótipo dos híbridos. Se o fenótipo é mutante, as mutações são alelos do mesmo gene; se é tipo selvagem, não são alelos. Neste caso, devemos portanto cruzar camundongo *little eye* com *tiny eye* e observar sua prole. Se a prole tiver olhos pequenos, as duas mutações são alelos do mesmo gene. Se tiverem olhos de tamanho normal, as duas mutações são alelos de genes diferentes. Para uma mutação dominante tal como *Eyeless*, não é possível nenhum teste de alelismo. Assim, não podemos determinar se *Eyeless* é um alelo da mutação *little eye* ou de *tiny eye*.

4. Distinga penetrância incompleta e expressividade variável.

Resposta: Penetrância incompleta ocorre quando um indivíduo com o genótipo para uma característica não expressa a característica desejada. Expressividade variável ocorre quando uma característica se manifesta em graus diferentes em um conjunto de indivíduos com o genótipo para a característica.

5. Em uma espécie de mosca, a cor de olho tipo selvagem é vermelha. Em uma linhagem mutante homozigota para a mutação w , a cor de olho é branco-pura; em outra linhagem mutante homozigota para a mutação y , a cor de olho é amarela. Os mutantes homozigotos brancos foram cruzados com mutantes homozigotos amarelos, e a prole toda tinha olhos vermelhos. Quando estas linhagens foram entrecruzadas, elas produziram três classes de prole: 92 de cor vermelha, 33 amarela e 41 branco-pura. (a) Pelos resultados destes cruzamentos, quantos genes controlam a cor de olho? Explique. (b) Se a resposta ao item (a) for maior que um, um gene é epistático em relação ao outro?

Resposta: Para responder (a), notamos que as moscas de F_1 todas tinham olhos vermelhos, isto é, tipo selvagem. As mutações w e y não são portanto alelos do mesmo gene, e concluímos que pelo menos dois genes devem controlar a cor de olho nesta espécie. Para responder (b), notamos que nas moscas F_2 , a proporção de segregação fenotípica desvia-se da 9:3:3:1 esperada para dois genes que segregam independentemente. A F_2 consiste em apenas três classes que, além disso, aparecem na proporção de cores 9 vermelhas:4 brancas:3 amarelas. Evidentemente, os homozigotos ww fazem com que as moscas tenham olhos brancos independente de quais alelos do gene y estejam presentes. Portanto, o gene w deve ser considerado epistático em relação ao gene y .

Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. Um geneticista obteve duas linhagens puras de camundongos, cada uma homozigota para uma mutação recessiva descoberta independentemente que impede a formação de pêlos no corpo. Uma linhagem mutante é chamada *naked* e a outra é chamada *hairless*. Para determinar se as duas mutações são alélicas, o geneticista cruzou camundongos *naked* e *hairless*. Toda a prole é fenotipicamente tipo selvagem; isto é, eles têm pêlos por todo o corpo. Após entrecruzar esses camundongos F_1 , o geneticista observou 115 tipo selvagem e 85 mutantes na F_2 . As mutações *naked* e *hairless* são alélicas? Como você explica a segregação de camundongos tipo selvagem e mutantes na F_2 ?

Resposta: As mutações *naked* e *hairless* não são alélicas porque os híbridos de F_1 são fenotipicamente tipo selvagem. Assim, as mutações *naked* e *hairless* são mutações de dois genes diferentes. Para explicar a proporção fenotípica na F_2 , primeiro adotaremos símbolos para tais mutações e seus alelos tipo selvagem dominantes:

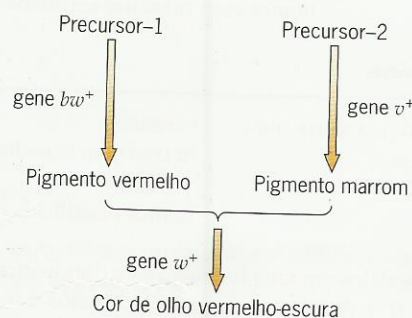
n = mutação *naked*, N = alelo tipo selvagem
 h = mutação *hairless*, H = alelo tipo selvagem

Com estes símbolos, os genótipos das linhagens parentais puras são nn *HH* (*naked*) e NN *hh* (*hairless*). Os híbridos de F_1 produzidos pelo cruzamento destas linhagens são, portanto, Nn *Hh*. Quando tais híbridos são entrecruzados, esperamos que surjam muitos genótipos diferentes na prole. Entretanto, cada alelo recessivo, quando homozigoto, impede a formação de pêlos no corpo. Assim, apenas camundongos que são genotipicamente $N-$ $H-$ desenvolverão pêlos; todos os outros – homozigotos nn ou homozigotos hh ou homozigotos para ambos os alelos recessivos – não desenvolverão pêlos no corpo. Podemos prever as freqüências dos fenótipos tipo selvagem e mutante caso seja suposto que os alelos *naked* e *hairless* segregam independentemente. A freqüência de camundongos que serão $N-$ $H-$ é $(3/4) \times (3/4) = 9/16 = 0,56$ (pela Regra de Multiplicação de Probabilidades) e a freqüência de camundongos que serão ou nn ou hh (ou ambos) é $(1/4) + (1/4) - [(1/4) \times (1/4)] = 7/16 = 0,44$ (pela Regra da Soma das Probabilidades). Assim, em uma amostra de 200 da prole F_2 , esperamos $200 \times 0,56 = 112$ sendo selvagens e $200 \times 0,44 = 88$ sendo mutantes. As freqüências observadas de 115 camundongos tipo selvagem e 85 mutantes estão próximas dos números esperados, sugerindo que a hipótese de dois genes

que segregam independentemente genes para pêlos no corpo é de fato correta.

2. Nas moscas-das-frutas, uma mutação recessiva, w , faz com que os olhos sejam brancos e outra mutação recessiva, v , faz com que sejam *vermilion*; e uma terceira mutação recessiva, bw , faz com que sejam de cor marrom. A cor de olho tipo selvagem é vermelho-escuro. Os híbridos produzidos pelo cruzamento de quaisquer dois mutantes homozigotos têm olhos escuros, e todas as combinações de mutantes duplos homozigotos têm olhos brancos. Quantos genes estas três mutações definem? Se a cor vermelho-escuro dos olhos tipo selvagem se deve ao acúmulo de dois pigmentos diferentes, um vermelho e outro marrom, qual gene controla a expressão de qual pigmento? Os genes podem ser ordenados em uma via para acúmulo de pigmento?

Resposta: As três mutações definem três genes diferentes porque, quando quaisquer duas mutações homozigotas são cruzadas, a prole tem cor de olho tipo selvagem. A mutação w impede a expressão de todo o pigmento porque as moscas homozigotas para ela não têm nem pigmento vermelho nem marrom em seus olhos; a mutação v impede a expressão do pigmento marrom, porque as moscas homozigotas para ele têm olhos *vermilion* (vermelho-claros); e a mutação bw impede a expressão do pigmento vermelho porque as moscas que são homozigotas para ele têm olhos marrom. Assim, o gene do tipo selvagem v controla a expressão do pigmento marrom, o gene do tipo selvagem bw controla a expressão do pigmento vermelho e o gene do tipo selvagem w é necessário para a expressão de ambos os pigmentos. Podemos resumir estes achados propondo que cada pigmento é expresso em uma via diferente e que o funcionamento destas vias depende do gene w do tipo selvagem.



Questões e Problemas Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 4.1. Que tipos sanguíneos podem ser observados nos filhos nascidos de uma mulher que tem tipo sanguíneo M e um homem com tipo sanguíneo MN?
- 4.2. Em coelhos, a coloração da pelagem depende dos alelos do gene c . Pela informação dada no capítulo, que fenótipos e proporções podem ser esperados dos seguintes cruzamentos: (a) $c^+c^+ \times cc$; (b) $c^+c \times c^+c$; (c) $c^+c^b \times c^+c^b$; (d) $cc^{cb} \times cc$; (e) $c^+c^b \times c^+c$; (f) $c^b c \times cc$?
- 4.3. Em camundongos, uma série de cinco alelos determina a cor da pelagem. Em ordem de dominância, estes alelos são: A^Y , pelagem amarela mas letal homozigoto; A^L , agouti com barriga clara; A^+ , agouti (tipo selvagem); a^+ , preto e castanho; e a , preto. Para cada um dos seguintes cruzamentos, cite a cor da pelagem dos genitores e as proporções fenotípicas esperadas entre a prole: (a) $A^Y A^L \times A^Y A^L$; (b) $A^Y a \times A^L a^+$; (c) $a^+ a \times A^Y a$; (d) $A^L a^+ \times A^L A^L$; (e) $A^L A^L \times A^Y A^+$; (f) $A^+ a^+ \times a^+ a$; (g) $a^+ a \times aa$; (h) $A^Y A^L \times A^+ a^+$; e (i) $A^Y a^+ \times A^Y A^+$.

4.4. Em várias plantas, tais como tabaco, primula e trevo vermelho, foi observado que combinações de alelos em ovócitos e pólen influenciam a compatibilidade reprodutiva das plantas. Combinações homozigotas, tais como S^1S^1 , não se desenvolvem porque o pólen S^1 não é efetivo em estigmas S^1 . Entretanto, o pólen S^1 é efetivo em estigmas S^2S^3 . Que prole pode ser esperada dos seguintes cruzamentos (semente genitora escrita primeiro): (a) $S^1S^2 \times S^2S^3$; (b) $S^1S^2 \times S^3S^4$; (c) $S^4S^5 \times S^4S^5$; e (d) $S^3S^4 \times S^5S^6$?

4.5. Das informações no capítulo sobre os tipos sanguíneos ABO, que fenótipos e proporções são esperados das seguintes reproduções: (a) $I^A I^A \times I^B I^B$; (b) $I^A I^B \times ii$; (c) $I^A i \times I^B i$; e (d) $I^A i \times ii$?

4.6. Uma mulher com tipo sanguíneo O teve uma criança, também do grupo O. A mulher diz que um homem tipo AB é o pai da criança. Sua afirmativa é verdadeira?

4.7. Outra mulher com tipo sanguíneo AB teve uma criança com tipo B. Dois homens diferentes se dizem o pai. Um tem sangue tipo A e o outro sangue tipo B. A evidência genética pode decidir em favor de um deles?

4.8. As cores das flores de plantas em determinada população podem ser azul, púrpura, turquesa, azul-claro ou branca. Uma série de cruzamentos entre membros diferentes da população produziram os seguintes resultados:

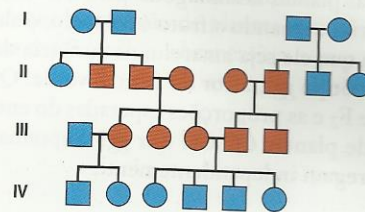
Cruzamento	Genitores	Prole
1	púrpura \times azul	toda púrpura
2	púrpura \times púrpura	76 púrpura, 25 turquesa
3	azul \times azul	86 azul, 29 turquesa
4	púrpura \times turquesa	49 púrpura, 52 turquesa
5	púrpura \times púrpura	69 púrpura, 22 azul
6	púrpura \times azul	50 púrpura, 51 azul
7	púrpura \times azul	54 púrpura, 26 azul, 25 turquesa
8	turquesa \times turquesa	toda turquesa
9	púrpura \times azul	49 púrpura, 25 azul, 23 azul-claro
10	azul-claro \times azul-claro	60 azul-claro, 29 turquesa, 31 branca
11	turquesa \times branca	toda azul-claro
12	branca \times branca	toda branca
13	púrpura \times branca	toda púrpura

Quanto genes e alelos estão envolvidos na herança de cor da flor? Indique todos os genótipos possíveis para os seguintes fenótipos: (a) púrpura; (b) azul; (c) turquesa; (d) azul-claro; (e) branco.

4.9. Uma mulher que tem sangue tipo O e tipo M casou com um homem que tem tipo AB e tipo MN. Se for suposto que os genes para os sistemas sanguíneos ABO e MN segregam independentemente, que tipos sanguíneos terão os filhos deste casal e em que proporções?

4.10. Uma linhagem japonesa de camundongos tem um andar peculiar descoordenado chamado valsador, que se deve a um alelo recessivo, v . O alelo dominante V faz com que o camundongo ande de modo coordenado. Um geneticista de camundongos recentemente isolou outra mutação recessiva que causa movimento descoordenado. Essa mutação, chamada *tango*, poderia ser um alelo do gene *waltzing* ou poderia ser uma mutação em um gene totalmente diferente. Proponha um teste para determinar se as mutações *waltzing* e *tango* são alelos e, se forem, proponha símbolos para indicá-los.

4.11. A surdez congênita em seres humanos é herdada como uma condição recessiva. No heredograma adiante, duas pessoas surdas, cada uma supostamente homozigota para uma mutação recessiva, casaram-se e tiveram quatro filhos com audição normal. Proponha uma explicação.



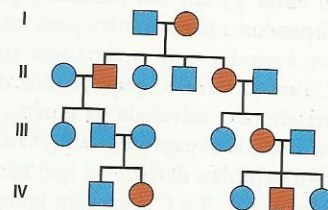
4.12. Na mosca-das-frutas, mutações recessivas em dois genes que segregam independentemente, *brown* e *purple*, impedem a síntese de pigmento vermelho nos olhos. Portanto, homozigotos para ambas estas mutações têm olhos marrom-púrpura. Entretanto, heterozigotos para ambas estas mutações têm olhos vermelho-escuros, isto é, olhos tipo selvagem. Se tais duplos heterozigotos forem entrecruzados, que tipos de prole serão produzidas e em que proporções?

4.13. A mutação dominante *Plum* na mosca-das-frutas também causa olhos marrom-púrpura. É possível determinar por experimentos genéticos se *Plum* é um alelo dos genes *brown* ou *purple*?

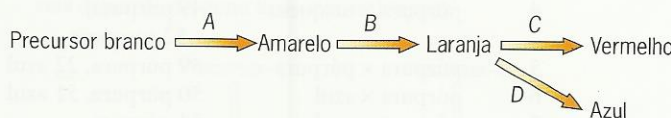
4.14. A partir de informação dada no capítulo, explique por que camundongos com pelagem amarela não são puros.

4.15. Um casal tem quatro filhos. Nem o pai nem a mãe exibe calvície; um dos dois filhos é calvo, mas nenhuma das filhas é calva.
(a) Se uma das filhas se casar com um homem não calvo e tiverem um filho, qual a chance de que o filho se torne calvo quando adulto?
(b) Se o casal tiver uma filha, qual a chance de que ela fique calva quando for adulta?

4.16. O heredograma a seguir mostra a herança de ataxia, um raro distúrbio neurológico caracterizado por movimentos descoordenados. A ataxia é causada por um alelo dominante ou recessivo? Explique.



- 4.17. As galinhas que possuem ambos os alelos para crista *rose* (*R*) e *pea* (*P*) têm cristas *walnut*, enquanto as galinhas que não têm nenhum destes alelos (isto é, são genotipicamente *rr pp*) têm cristas *single*. Pelas informações sobre interações desses dois genes dadas no capítulo, determine os fenótipos e as proporções esperadas dos seguintes cruzamentos:
- $RR Pp \times rr Pp$;
 - $rr PP \times Rr Pp$;
 - $Rr Pp \times Rr pp$;
 - $Rr pp \times rr pp$.
- 4.18. Galinhas de crista *rose* cruzadas com galinhas de crista tipo *walnut* produziram 15 galinhas *walnut*, 14 *rose*, 5 *pea* e 6 com cristas *single*. Determine os genótipos dos genitores.
- 4.19. As abóboras com o alelo dominante *C* têm um fruto branco, enquanto as plantas homozigotas para o alelo recessivo *c* têm fruto colorido. Quando o fruto é colorido, o alelo dominante *G* faz com que ele seja amarelo; na ausência deste alelo (isto é, com genótipo *gg*), a cor do fruto é verde. Quais são os fenótipos de F_2 e as proporções esperadas do entrecruzamento da prole de plantas $CC GG$ e $cc gg$? Suponha que os genes *C* e *G* segregam independentemente.
- 4.20. A raça branca Leghorn de galinhas é homozigota para o alelo dominante *C*, que produz plumagem colorida. Entretanto, esta raça também é homozigota para o alelo dominante *I* de um gene que segrega independentemente e inibe a coloração da plumagem. Consequentemente, as galinhas têm plumagem branca. A raça branca Wyandotte de galinhas não tem nem o alelo para cor nem o inibidor de cor. Ela é, portanto, genotipicamente $cc ii$. Quais são os fenótipos de F_2 e as proporções esperadas do entrecruzamento da prole de uma fêmea branca Leghorn e um galo branco Wyandotte?
- 4.21. *Drosophila* homozigotas para a mutação recessiva *scarlet* têm olhos vermelho-claros porque não podem produzir o pigmento marrom. As *Drosophila* homozigotas para a mutação recessiva *brown* têm olhos marrom-púrpura porque não podem sintetizar pigmento vermelho. *Drosophila* homozigotas para ambas estas mutações têm olhos brancos porque não podem produzir nenhum dos tipos de pigmento. As mutações *brown* e *scarlet* segregam independentemente. Se *Drosophila* que são heterozigotas para ambas estas mutações são entrecruzadas, que tipo de prole elas produzirão e em que proporções?
- 4.22. Considere o seguinte esquema hipotético de determinação da cor da pelagem em um mamífero. O gene *A* controla a conversão de um pigmento branco P_0 em um pigmento cinza P_1 ; o alelo dominante *A* produz a enzima necessária para esta conversão e o alelo recessivo *a* produz uma enzima sem atividade bioquímica. O gene *B* controla a conversão do pigmento cinza P_1 em um pigmento preto P_2 ; o alelo dominante *B* produz a enzima ativa para esta conversão e o alelo recessivo *b* produz uma enzima sem atividade. O alelo dominante *C* de um terceiro gene produz um polipeptídeo que inibe totalmente a atividade da enzima produzida pelo gene *A*; isto é, impede a reação $P_0 \rightarrow P_1$. O alelo *c* deste gene produz um polipeptídeo defeituoso que não inibe a reação $P_0 \rightarrow P_1$. Os genes *A*, *B* e *C* segregam independentemente, e nenhum outro gene está envolvido. Na F_2 do cruzamento $AA bb CC \times aa BB cc$, qual a proporção de segregação fenotípica esperada?
- 4.23. Que proporção de segregação fenotípica de F_2 seria esperada no cruzamento descrito no problema anterior se o alelo dominante, *C*, do terceiro gene fizesse um produto que inibisse totalmente a atividade da enzima produzida pelo gene *B*, isto é, impedisse a reação $P_1 \rightarrow P_2$, em vez de inibir a atividade da enzima produzida pelo gene *A*?
- 4.24. O martim-pescador da micronésia, *Halcyon cinnamomina*, tem uma face de cor de canela. Em algumas aves, a cor continua até o tórax, produzindo um dentre três padrões: um círculo, um escudo ou um triângulo; em outras aves, não há cor no tórax. Um macho com um triângulo colorido foi cruzado com uma fêmea que não tem cor no tórax, e toda a sua prole tinha um escudo colorido no tórax. Quando estas proles foram entrecruzadas, elas produziram uma F_2 com uma proporção fenotípica de 3 círculo:6 escudo:3 triângulo:4 sem cor. (a) Determine o modo de herança para esta característica e indique os genótipos das aves em todas as três gerações. (b) Se um macho sem cor no tórax é cruzado com uma fêmea com um escudo colorido no tórax e a F_1 segrega em uma proporção de 1 círculo:2 escudo:1 triângulo, quais são os genótipos dos genitores e sua prole?
- 4.25. Em uma espécie de árvore, a cor da semente é determinada por quatro genes que segregam independentemente: *A*, *B*, *C* e *D*. Os alelos recessivos de cada um destes genes (*a*, *b*, *c* e *d*) produzem enzimas anormais que não podem catalisar uma reação na via biossintética para o pigmento da semente. Esta via é diagramada como a seguir:



Quando tanto os pigmentos vermelho e azul estão presentes, as sementes são púrpura. Árvores com os genótipos $Aa Bb Cc Dd$ e $Aa Bb Cc dd$ foram cruzadas.

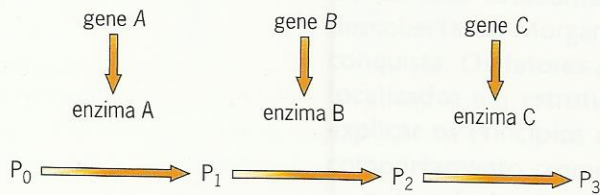
- Que cor têm as sementes nestes dois genótipos parentais?
- Que proporção da prole do cruzamento terá sementes brancas?
- Determine as proporções relativas de prole vermelha, branca e azul do cruzamento.

- 4.26. Foram feitos vários cruzamentos entre linhagens puras de labradores pretos e amarelos. Toda a prole F_1 era preta. Quando esta prole foi entrecruzada, foi produzida uma F_2 que consistia em 91 pretos, 39 amarelos e 30 chocolate. (a) Proponha uma explicação para a herança da cor da pelagem em labradores. (b) Proponha uma via bioquímica para a determinação da cor de pelagem e indique como os genes relevantes controlam a cor da pelagem.

- 4.27. Duas plantas com flores brancas, cada uma de linhagens puras, foram cruzadas. Todas as plantas F_1 tinham flores vermelhas. Quando estas plantas F_1 foram entrecruzadas, elas

produziram uma F_2 que consistia em 177 plantas com flores vermelhas e 142 com flores brancas. (a) Proponha uma explicação para a herança da cor da flor nesta espécie de planta. (b) Proponha uma via bioquímica para a pigmentação da flor e indique quais genes controlam quais etapas nesta via.

- 4.28. Considere a seguinte via biossintética geneticamente controlada para pigmentos nas flores de uma planta hipotética:



Suponha que o gene *A* controla a conversão de um pigmento branco, P_0 , em outro pigmento branco, P_1 ; o alelo dominante

A especifica uma enzima necessária para essa conversão e o alelo recessivo *a* especifica uma enzima defeituosa sem função bioquímica. O gene *B* controla a conversão do pigmento branco, P_1 , em um pigmento rosa, P_2 ; o alelo dominante, *B*, produz a enzima necessária para esta conversão e o alelo recessivo, *b*, produz uma enzima defeituosa. O alelo dominante, *C*, do terceiro gene especifica uma enzima que converte o pigmento rosa, P_2 , em um pigmento vermelho, P_3 ; seu alelo recessivo, *c*, produz uma enzima alterada que não pode efetuar esta conversão. O alelo dominante, *D*, de um quarto gene produz um polipeptídeo que inibe totalmente a função da enzima *C*; isto é, bloqueia a reação $P_2 \rightarrow P_3$. Seu alelo recessivo, *d*, produz um polipeptídeo defeituoso que não bloqueia esta reação. Suponha que a cor da flor seja determinada apenas por estes quatro genes e que eles segregam independentemente. Na F_2 de um cruzamento entre plantas com o genótipo *AA bb CC DD* e plantas com o genótipo *aa BB cc dd*, que proporção de plantas terá (a) flores vermelhas? (b) flores rosa? (c) flores brancas?

A Base Cromossômica do Mendelismo

Tópicos

- Cromossomos
- A Teoria Cromossômica da Hereditariedade
- Genes Ligados ao Sexo em Seres Humanos
- Cromossomos Sexuais e Determinação do Sexo
- Compensação de Dose de Genes Ligados ao X

Sexo, Cromossomos e Genes

O que faz com que os organismos se desenvolvam como machos ou fêmeas? Por que existem apenas dois fenótipos sexuais? O sexo de um organismo é determinado por seus genes? Estas e perguntas correlatas têm intrigado geneticistas desde a redescoberta do trabalho de Mendel no começo do século vinte.

A descoberta de que os genes têm um papel na determinação do sexo surgiu de uma fusão entre duas disciplinas antes cientificamente distintas, a genética – o estudo da hereditariedade – e a citologia – o estudo das células. No começo do século vinte, estas disciplinas foram reunidas por uma amizade entre dois renomados cientistas americanos, Thomas Hunt Morgan e Edmund Beecher Wilson. Morgan era o geneticista e Wilson o citologista.

Como citologista, Wilson estava interessado no comportamento dos cromossomos. Tais estruturas demonstram-se importantes para a determinação do sexo em muitas espécies, incluindo a nossa. Wilson foi um dos primeiros a investigar as diferenças nos cromossomos dos dois sexos. Através de cuidadosos estudos, ele e



Macho e fêmea de *Drosophila*. O macho possui olhos brancos porque tem uma mutação em seu cromossomo X.

seus colegas mostraram que tais diferenças estavam confinadas a um par especial de cromossomos chamados cromossomos sexuais. Wilson descobriu que o comportamento destes cromossomos durante a meiose poderia explicar a herança do sexo.

Como geneticista, Morgan estava interessado na identificação de genes. Ele enfocou sua pesquisa na mosca-das-frutas, *Drosophila melanogaster*, e rapidamente descobriu um gene que dava proporções fenotípicas diferentes em machos e fêmeas. Morgan imaginou que esse gene estava situado em um dos cromossomos sexuais, e um de seus estudantes, Calvin Bridges, provou que sua hipótese estava correta. A descoberta de Morgan de que os genes residem em cromossomos foi uma grande conquista. Os fatores genéticos abstratos postulados por Mendel foram finalmente localizados em estruturas visíveis dentro das células. Os geneticistas agora podiam explicar os Princípios da Segregação e da Distribuição Independente em termos de comportamento cromossômico.

A descoberta de que genes específicos determinam o sexo de um organismo veio muito depois, apenas após uma outra disciplina científica, a biologia molecular, ter unido forças com a genética e a citologia. Através de seus esforços combinados, citologistas, geneticistas e biólogos moleculares identificaram genes determinantes de sexo ao estudar indivíduos raros nos quais o fenótipo sexual era inconsistente com os cromossomos sexuais que estavam presentes. Hoje em dia, pesquisadores em todos os três campos estão tentando explicar como tais genes controlam o desenvolvimento sexual.

CROMOSSOMOS

Cada espécie tem um conjunto característico de cromossomos.

Os cromossomos foram descobertos na segunda metade do século dezenove por um citologista alemão, W. Waldeyer. Pesquisas subseqüentes com muitos organismos diferentes estabeleceram que os cromossomos são característicos dos núcleos de todas as células. Eles são mais bem vistos com a aplicação de corantes às células em divisão. Durante a divisão, o material em um cromossomo é compactado em um pequeno volume, o que lhe confere o aspecto de um cilindro compacto organizado. Durante a interfase das divisões celulares, os cromossomos não são facilmente vistos, mesmo com os melhores corantes. Os cromossomos interfásicos são pouco helicoidizados, formando filamentos finos que são distribuídos pelo núcleo. Conseqüentemente, quando são aplicados corantes, todo o núcleo é corado e cromossomos individuais não podem ser identificados. Este emaranhado difuso de filamentos é chamado **cromatina**. Algumas regiões da cromatina coram-se mais escuro do que outras, sugerindo uma diferença subjacente de organização. As regiões claras e escuras, respectivamente chamadas **eucromatina** (do grego “verdadeiro”) e **heterocromatina** (da palavra grega para “diferente”), têm densidades diferentes de filamentos cromossômicos.

Número de Cromossomos

Dentro de uma espécie, o número de cromossomos é quase sempre um múltiplo inteiro de um número básico. Nos seres

humanos, por exemplo, o número básico é 23. Ovócitos e espermatozoides maduros têm este número de cromossomos. A maioria dos outros tipos de células humanas têm o dobro (46), embora algumas, tais como as células do fígado, têm quatro vezes (92) o número básico.

O número (**n**) cromossômico **haplóide**, ou básico, define um conjunto de cromossomos chamado **genoma haplóide**. A maioria das células somáticas contém dois de cada um dos cromossomos neste conjunto e são portanto **diplóides (2n)**. As células com quatro de cada um dos cromossomos são **tetraplóides (4n)**, as com oito são **octaplóides (8n)** e assim em diante.

O número básico de cromossomos varia entre espécies. O número de cromossomos não está relacionado com o tamanho ou a complexidade biológica de um organismo, a maioria das espécies contendo entre 10 e 40 cromossomos em seus genomas (Quadro 5.1). O muntjac, um cervo asiático, tem apenas três cromossomos em seu genoma, enquanto algumas espécies de samambaias têm muitas centenas.

Cromossomos Sexuais

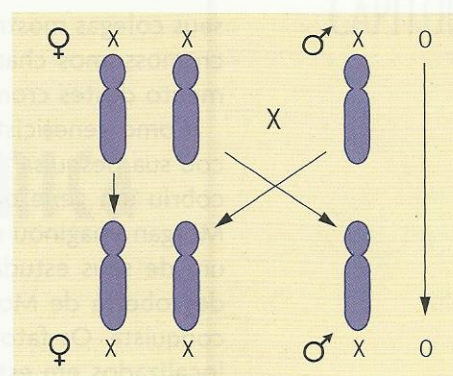
Algumas espécies animais, tais como fêmeas de gafanhotos, têm um cromossomo a mais que os machos (Fig. 5.1a). Esse cromossomo extra, originalmente observado em outros insetos, é chamado **cromossomo X**. Fêmeas destas espécies têm dois cromossomos X e machos têm apenas um. Assim, as fêmeas são citologicamente XX e os machos são X0, em que o “0” indica a falta de um cromossomo. Durante a meiose na fêmea, os dois cromossomos X formam um par

QUADRO 5.1

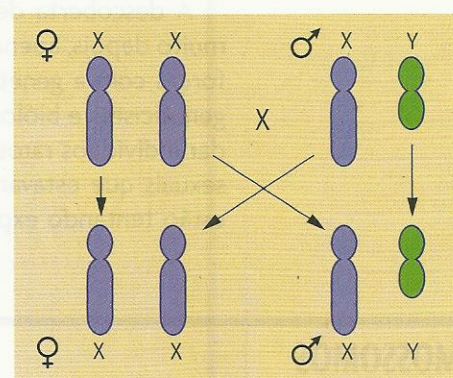
Número de Cromossomos em Diferentes Organismos	
Organismo	Nº Haplóide de Cromossomos
Eucariontes Simples	
Leveduras (<i>Saccharomyces cerevisiae</i>)	16
Fungo do pão (<i>Neurospora crassa</i>)	7
Alga verde unicelular (<i>Chlamydomonas reinhardtii</i>)	17
Plantas	
Milho (<i>Zea mays</i>)	10
Trigo do pão (<i>Triticum aestivum</i>)	21
Tomate (<i>Lycopersicon esculentum</i>)	12
Fava (<i>Vicia faba</i>)	6
Sequóia gigante (<i>Sequoia sempervirens</i>)	11
Crucífera (<i>Arabidopsis thaliana</i>)	5
Animais Invertebrados	
Mosca-das-frutas (<i>Drosophila melanogaster</i>)	4
Mosquito (<i>Anopheles culicifacies</i>)	3
Estrela-do-mar (<i>Asterias forbesi</i>)	18
Nematódeo (<i>Caenorhabditis elegans</i>)	6
Mexilhão (<i>Mytilus edulis</i>)	14
Animais Vertebrados	
Ser humano (<i>Homo sapiens</i>)	23
Chimpanzé (<i>Pan troglodytes</i>)	24
Gato (<i>Felis domesticus</i>)	36
Camundongo (<i>Mus musculus</i>)	20
Galinha (<i>Gallus domesticus</i>)	39
Sapo (<i>Xenopus laevis</i>)	17
Peixe (<i>Esox lucius</i>)	25

e então se separam, produzindo ovócitos que contêm um único cromossomo X. Durante a meiose no macho, o cromossomo X único move-se independentemente dos outros cromossomos e é incorporado à metade dos espermatozoides; a outra metade não recebe nenhum cromossomo X. Assim, quando espermatozoides e ovócitos se unem, são produzidos dois tipos de zigotos: XX, que se desenvolvem em fêmeas, e X0, que se desenvolvem em machos. Como cada um destes tipos é igualmente provável, o mecanismo reprodutivo preserva uma proporção de 1:1 de machos para fêmeas nestas espécies.

Em muitos outros animais, incluindo seres humanos, homens e mulheres têm o mesmo número de cromossomos (Fig. 5.1b). Esta igualdade numérica é devida à presença de um cromossomo no homem, chamado **cromossomo Y**, que faz par com o X durante a meiose. O cromossomo Y é morfológicamente distinguível do cromossomo X. Nos humanos, por exemplo, o Y é menor que o X, e seu centrômero está situado perto de uma das pontas (Fig. 5.2). O material comum aos cromossomos X e Y é limitado, consistindo principalmente em segmentos curtos perto das pontas dos cro-



(a) Herança de cromossomos sexuais em animais com fêmeas XX e machos X0.



(b) Herança de cromossomos sexuais em animais com fêmeas XX e machos XY.

Fig. 5.1 ■ Herança de cromossomos sexuais em animais.

mossomos. Durante a meiose no homem, os cromossomos X e Y separam-se um do outro, produzindo dois tipos de espermatozoides, contendo X e contendo Y. As frequências dos dois tipos são aproximadamente iguais. As mulheres XX produzem ovócitos apenas de um tipo, contendo X. Se a fertilização ocorrer aleatoriamente, cerca de metade dos zigotos será XX e metade será XY, levando a uma proporção de concepção de 1:1. Entretanto, em seres humanos, os espermatozoides contendo Y têm uma vantagem de fertilização, e a proporção sexual zigótica é de cerca de 1,3:1. Durante o desenvolvimento, o excesso de homens é diminuído pela via-



Fig. 5.2 ■ Cromossomos humanos X e Y. As regiões terminais são comuns a ambos os cromossomos sexuais.

bilidade diferencial de embriões XX e XY, e, ao nascimento, homens são apenas um pouco mais numerosos que mulheres (proporção sexual 1,07:1). Na idade reprodutiva, o excesso de homens é essencialmente eliminado, e a proporção sexual é próxima de 1:1.

Os cromossomos X e Y são chamados **cromossomos sexuais**. Todos os outros cromossomos no genoma são chamados **autossomos**. Os cromossomos sexuais foram descobertos nos primeiros anos do século vinte com o trabalho dos citologistas americanos C. E. McClung, N. M. Stevens, W. S. Sutton, e E. B. Wilson. Essa descoberta coincidiu aproximadamente com o surgimento do mendelismo e estimulou pesquisas sobre as possíveis relações entre os princípios de Mendel e o comportamento meiótico dos cromossomos.

PONTOS IMPORTANTES

- Os cromossomos individuais tornam-se visíveis durante a divisão celular; entre as divisões, eles formam um emaranhado de fibras chamado cromatina.
- As células somáticas diplóides têm o dobro de cromossomos que os gametas haplóides.
- As células de homens e mulheres podem ter números diferentes de cromossomos X e Y; entretanto, o número de autossomos nestas células é o mesmo.

A TEORIA CROMOSSÔMICA DA HEREDITARIEDADE

Os estudos da herança de uma característica ligada ao sexo em *Drosophila* forneceram a evidência de que o comportamento meiótico dos cromossomos é a base dos Princípios da Segregação e da Distribuição Independente de Mendel.

Em 1910, muitos biólogos suspeitaram que os genes estivessem situados nos cromossomos, mas eles não tinham uma prova definitiva. Os pesquisadores precisavam encontrar um gene que pudesse ser inequivocamente ligado a um cromossomo. Esta meta requeria que o gene fosse definido por um alelo mutante e que o cromossomo fosse morfológicamente distinguível. Além disso, o padrão de transmissão gênica tinha que refletir o comportamento do cromossomo durante a reprodução. Todos estes requisitos foram atendidos quando o biólogo americano Thomas H. Morgan descobriu uma determinada mutação para cor de olho na mosca-das-frutas, *Drosophila melanogaster*. Morgan começou seus experimentos com esta espécie de mosca em 1909. Ela era idealmente adequada para as pesquisas genéticas porque se reproduz rapidamente e prolificamente e era de baixo custo para se criar em laboratório. Além disso, tem apenas quatro pares de cromossomos, sendo um par de cromossomos sexuais, XX nas fêmeas e XY nos machos. Os cromossomos X e Y foram morfológicamente distinguíveis um do outro e de cada um dos autossomos. Por

meio de cuidadosos experimentos, Morgan foi capaz de mostrar que a mutação de cor de olho era herdada junto com o cromossomo X, sugerindo que o gene para cor de olho estava fisicamente situado neste cromossomo. Mais tarde, um de seus estudantes, Calvin B. Bridges, obteve a prova definitiva de sua Teoria Cromossômica da Hereditariedade.

Evidência Experimental Ligando a Herança de Genes a Cromossomos

Os experimentos de Morgan começaram com sua descoberta de um macho mutante de drosófila que tinha olhos brancos em vez de vermelhos do tipo selvagem. Quando este macho era cruzado com fêmeas tipo selvagem, toda a prole tinha olhos vermelhos, indicando que branco é recessivo em relação a vermelho. Quando essa prole foi entrecruzada, Morgan observou um padrão peculiar de segregação: todas as filhas, mas apenas metade dos filhos, tinham olhos vermelhos; a outra metade dos filhos tinha olhos brancos. Este padrão sugeriu que a herança da cor de olho era ligada aos cromossomos sexuais. Morgan propôs que o gene para cor de olho estava presente no cromossomo X, mas não no Y, e que os fenótipos branco e vermelho eram devidos a dois alelos diferentes, um alelo mutante indicado por w e um alelo tipo selvagem indicado por w^+ .

A hipótese de Morgan é diagramada na Fig. 5.3. As fêmeas tipo selvagem no primeiro cruzamento supostamente

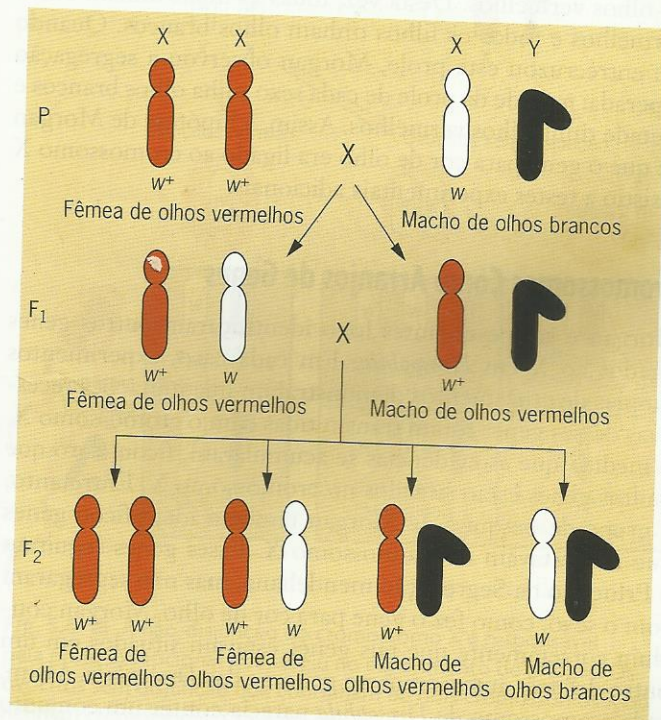


Fig. 5.3 ■ Experimento de Morgan sobre a herança de olhos brancos em *Drosophila*. A transmissão da condição mutante em associação ao sexo sugeriu que o gene para cor de olho estava presente no cromossomo X mas não no cromossomo Y.

são homozigotas para o alelo w^+ . Seu cruzamento supostamente leva o alelo mutante w em seu cromossomo X e nenhum dos alelos no cromossomo Y. Um organismo que tem apenas uma cópia de um gene é chamado **hemizigoto**. Entre a prole do cruzamento, os filhos herdam um cromossomo X de sua mãe e um cromossomo Y de seu pai. Como o X herdado da mãe leva o alelo w^+ , esses filhos têm olhos vermelhos. As filhas, em contraste, herdam um cromossomo X de cada um dos genitores, um X com w^+ da mãe e um X com w do pai. Entretanto, como w^+ é dominante em relação a w , essas fêmeas heterozigotas de F_1 também têm olhos vermelhos.

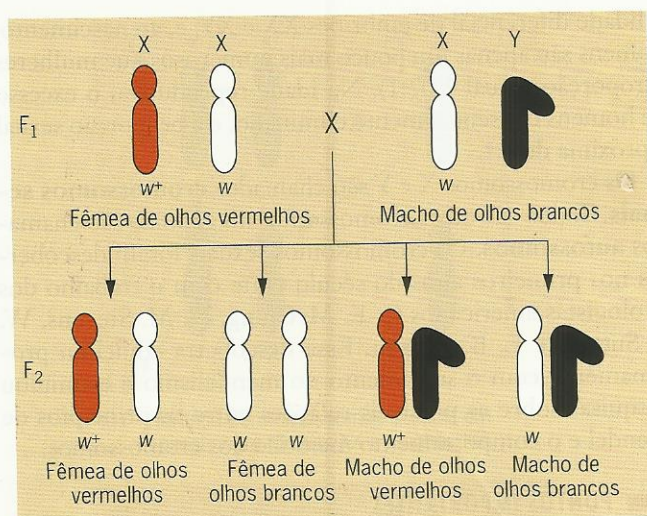
Quando os machos e fêmeas de F_1 são entrecruzados, quatro classes genóticas de prole são produzidas, cada uma representando uma combinação diferente de cromossomos sexuais. As moscas XX, que são fêmeas, têm olhos vermelhos porque pelo menos um alelo w^+ está presente. As moscas XY, que são machos, têm olhos vermelhos ou olhos brancos, dependendo de qual cromossomo X é herdado das fêmeas heterozigotas de F_1 . A segregação dos alelos w e w^+ nestas fêmeas é portanto o motivo pelo qual metade dos machos de F_2 têm olhos brancos.

Morgan fez experimentos adicionais para confirmar os elementos de sua hipótese. Em um deles (Fig. 5.4a), ele cruzou fêmeas de F_1 supostamente heterozigotas para o gene de cor de olho com machos brancos mutantes. Como ele esperava, metade da prole de cada sexo tinha olhos brancos e metade tinha olhos vermelhos. Em outro experimento (Fig. 5.4b), ele cruzou fêmeas com olhos brancos com machos de olhos vermelhos. Desta vez, todas as filhas tinham olhos vermelhos e todos os filhos tinham olhos brancos. Quando ele entrecruzou esta prole, Morgan observou a segregação esperada: metade da prole de cada sexo tinha olhos brancos e metade tinha olhos vermelhos. Assim, a hipótese de Morgan de que o gene para cor de olho era ligado ao cromossomo X resistiu a testes experimentais adicionais.

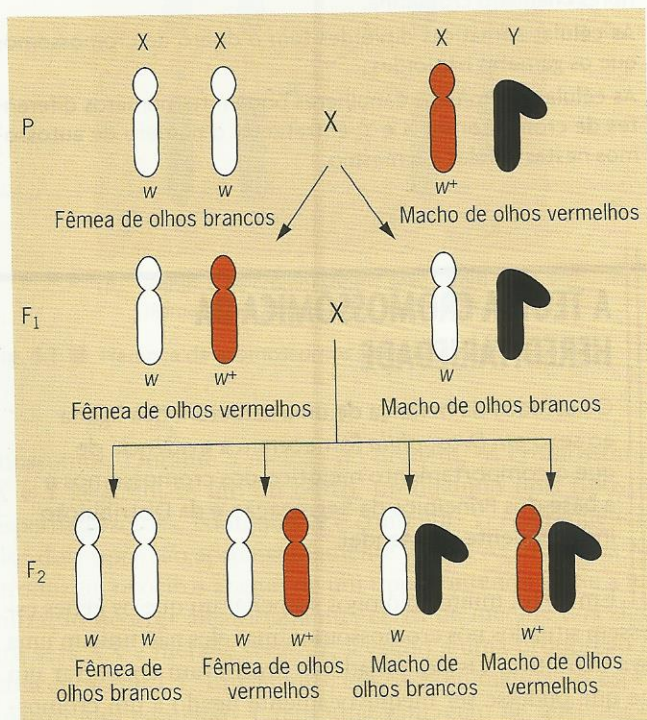
Cromossomos Como Arranjos de Genes

Morgan e seus estudantes logo identificaram outros genes ligados ao X em *Drosophila*. Em cada caso, experimentos simples de cruzamentos demonstraram que mutações recessivas desses genes foram transmitidas com o cromossomo X. À medida que as evidências se acumularam, ficou claro que muitos genes estão situados no cromossomo X. Entretanto, o grupo de pesquisa de Morgan também identificou genes que não estavam no cromossomo X. Esses genes seguiram o Princípio da Segregação mendeliana, mas não segregaram com o sexo, como fez o gene para cor de olho. Morgan concluiu corretamente que tais genes estavam situados em um outro dos três autossomos no genoma das *Drosophila*. Assim, cada cromossomo de *Drosophila* parecia conter um conjunto diferente de genes.

O laboratório de Morgan então tentou determinar as relações entre os genes em um determinado cromossomo. Eles continuaram com a suposição de que os genes eram dispos-



(a) Cruzamento entre uma fêmea heterozigota e um macho mutante hemizigoto.



(b) Cruzamento entre uma fêmea mutante homozigota e um macho tipo selvagem hemizigoto.

Fig. 5.4 ■ Testes experimentais da hipótese de Morgan de que o gene para cor de olho em *Drosophila* é ligado ao X. Em cada experimento, a cor de olho é herdada com o cromossomo X. Assim, os resultados desses cruzamentos apoiaram a hipótese de Morgan de que o gene para cor de olho é ligado ao X.

tos de modo linear, uma idéia inspirada na evidência citológica de que o cromossomo era um filamento longo e fino. Em apenas alguns anos, os estudos de Morgan foram capazes de mostrar que os genes de fato estavam situados em locais diferentes, ou *loci* (da palavra latina para “lugar”; singular, *locus*), em uma estrutura linear. Esta análise, que discutiremos

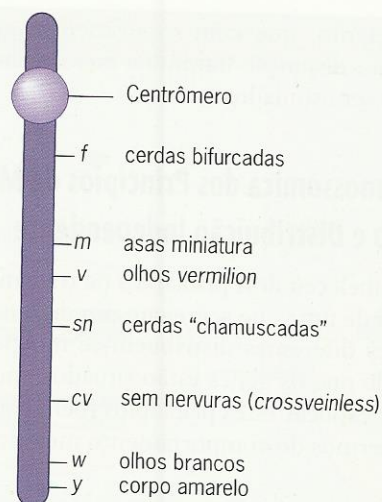


Fig. 5.5 ■ Um mapa de genes no cromossomo X de *Drosophila*.

no Cap. 7, produziu os primeiros mapas genéticos do mundo – diagramas mostrando as posições dos genes e as relativas distâncias entre eles (Fig. 5.5). O laboratório de Morgan foi pioneiro nos métodos para mapeamento genético e criou a fundação para pesquisas subseqüentes sobre a estrutura física dos cromossomos. Finalmente, a linearidade dos cromossomos foi associada à estrutura linear do DNA (veja Cap. 9).

Estes estudos iniciais com *Drosophila*, primariamente o trabalho de Morgan e seus estudantes (veja Um Marco na Genética: A Sala das Moscas de Morgan, no final deste capítulo), fortaleceram muito a visão de que todos os genes estão localizados em cromossomos e que os experimentos de Mendel podiam ser explicados pelas propriedades de transmissão dos cromossomos durante a reprodução. Esta idéia, chamada **Teoria Cromossômica da Hereditariedade**, é uma das mais importantes conquistas da biologia. Desde sua formulação na primeira parte do século vinte, a Teoria Cromossômica da Hereditariedade forneceu um arcabouço unificante para todos os estudos de herança.

Não-disjunção Como Prova da Teoria Cromossômica

Morgan mostrou que um gene da cor de olho estava no cromossomo X de drosófilas correlacionando a herança deste gene com a transmissão do cromossomo X durante a reprodução. Entretanto, como já foi notado, foi um de seus estudantes, C. B. Bridges, que provou a teoria comossômica mostrando que exceções às regras também podiam ser explicadas por comportamento cromossômico.

Bridges fez um dos experimentos de Morgan em uma escala maior. Ele cruzou fêmeas de drosófilas de olhos brancos com machos de olhos vermelhos e examinou muitas proles F_1 . Embora, como esperado, quase todas as moscas F_1 fossem ou fêmeas de olhos vermelhos ou machos de olhos brancos, Bridges descobriu algumas moscas excepcionais – fêmeas de olhos brancos e machos de olhos vermelhos. Ele

cruzou essas exceções para determinar como elas podem ter surgido. Os machos excepcionais eram todos estéreis, porém as fêmeas excepcionais eram férteis, e, quando cruzadas com machos normais de olhos vermelhos, produziram muita prole, incluindo grande número de filhas de olhos brancos e filhos de olhos vermelhos. Assim, as fêmeas excepcionais da F_1 , embora raras, produziam muita prole excepcional.

Bridges explicou estes resultados propondo que as moscas excepcionais de F_1 eram o resultado de comportamento anormal do cromossomo X durante a meiose em fêmeas da geração P. Geralmente, os cromossomos X nestas fêmeas devem sofrer **disjunção**, ou se separarem, durante a meiose. Entretanto, ocasionalmente, eles podem falhar em se separar, produzindo um ovócito com dois cromossomos X ou um ovócito sem nenhum cromossomo X. A fertilização de tais ovócitos anormais por espermatozóides normais produziria zigotos com um número anormal de cromossomos sexuais. A Fig. 5.6 ilustra as possibilidades.

Se um ovócito com dois cromossomos X (geralmente chamados de ovócitos duplo-X; genótipo X^wX^w) é fertilizado

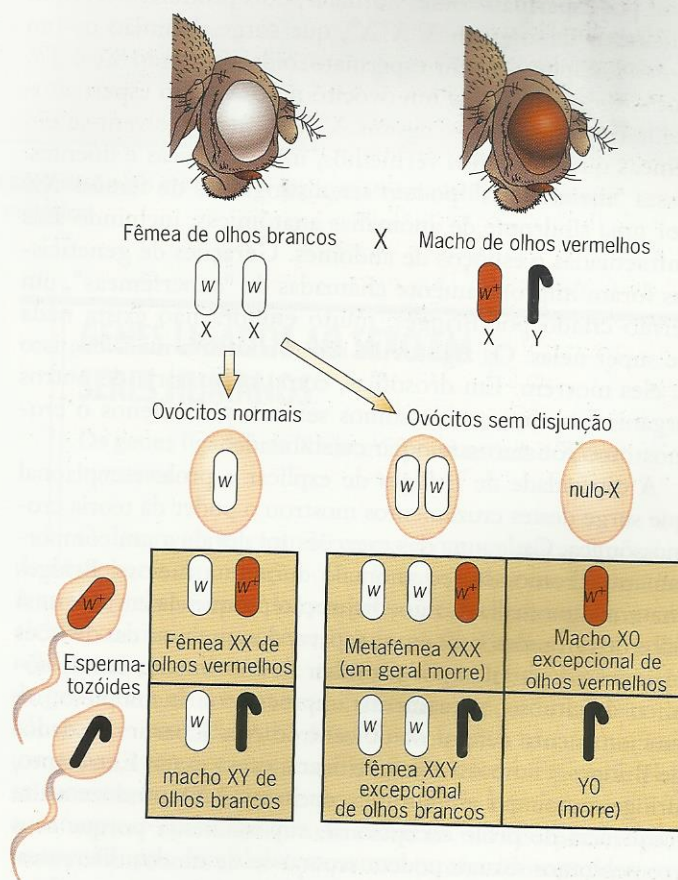


Fig. 5.6 ■ A não-disjunção do cromossomo X é responsável pela prole excepcional que surgiu no experimento de Bridges. Os ovócitos sem disjunção que continham ou dois cromossomos X ou nenhum cromossomo X uniram-se a espermatozóides normais que continham ou um X ou um Y para produzir quatro tipos de zigotos. Os zigotos XXY desenvolveram-se em fêmeas de olho branco, os zigotos XO desenvolveram-se em machos estéréis de olho vermelho e os zigotos XXX e YO morreram.

por um espermatozoide contendo Y, o zigoto será X^wX^wY . Como cada um dos cromossomos X neste zigoto leva um alelo mutante w , a mosca resultante terá olhos brancos. Se um ovócito sem um cromossomo X (geralmente chamado ovócito X nulo) é fertilizado por um espermatozoide portador de X (X^+), o zigoto será X^+0 . (Novamente, "0" indica a ausência de um cromossomo.) Como o único X neste zigoto leva um alelo w^+ , o zigoto se desenvolverá em uma mosca de olhos vermelhos. Bridges deduziu que as moscas XXY eram fêmeas e as moscas X0 eram machos. As fêmeas excepcionais de olhos brancos que ele observou eram portanto X^wX^wY e os machos excepcionais de olhos vermelhos eram X^+0 . Bridges confirmou as constituições cromossômicas destas moscas excepcionais por observação citológica direta. Como os animais X0 eram machos, Bridges concluiu que em drosófilas o cromossomo Y não guardava nenhuma relação com a determinação do fenótipo sexual. Entretanto, como os machos X0 eram sempre estéreis, ele percebeu que este cromossomo deveria ser importante para funcionamento sexual masculino.

Bridges reconheceu que a fertilização de ovócitos anormais por espermatozoides normais podia produzir dois tipos adicionais de zigotos: $X^wX^wX^+$, que surge da união de um ovócito duplo-X e um espermatozoide contendo X, e $Y0$, que surge da união de um ovócito nulo-X e um espermatozoide contendo Y. Os zigotos $X^wX^wX^+$ desenvolvem-se em fêmeas que têm olhos vermelhos, mas são fracas e doentes. Essas "metafêmeas" podem ser distinguidas de fêmeas XX por uma síndrome de anomalias anatômicas, incluindo asas anfractuadas e esboços de abdômes. Gerações de geneticistas foram impropriamente chamadas de "superfêmeas", um termo criado por Bridges, muito embora não exista nada de super nelas. Os zigotos $Y0$ são totalmente inviáveis; isto é, eles morrem. Em drosófilas, como na maioria de outros organismos com cromossomos sexuais, pelo menos o cromossomo X é necessário para viabilidade.

A capacidade de Bridges de explicar a prole excepcional que surge destes cruzamentos mostrou o poder da teoria cromossômica. Cada uma das exceções foi devida a um comportamento cromossômico anômalo durante a meiose. Bridges chamou a anomalia de **não-disjunção** porque ela envolve uma falha dos cromossomos em se separar durante uma das divisões meióticas. Essa falha pode resultar de movimento cromossômico defeituoso, pareamento impreciso ou incompleto, ou funcionamento defeituoso do centrômero. A partir dos dados de Bridges, é impossível especificar a causa exata. Entretanto, Bridges notou que as fêmeas excepcionais XXY produzem alta frequência de prole excepcional, supostamente porque seus cromossomos sexuais podem separar-se de modos diferentes: os cromossomos X podem desunir-se um do outro, ou o X separar-se do Y. Neste último caso, um ovócito duplo- ou nulo-X é produzido porque o X que não se separa do Y está livre para se mover para um dos pólos durante a primeira divisão meiótica. Quando fertilizados por um espermatozoide normal, esses ovócitos anormais produzirão zigotos excepcionais.

Bridges observou os efeitos da não-disjunção cromossômica que ocorreu durante a meiose em fêmeas. Devemos

notar, entretanto, que com experimentos apropriados os efeitos da não-disjunção durante a meiose em machos também podem ser estudados.

A Base Cromossômica dos Princípios de Mendel de Segregação e Distribuição Independente

Mendel estabeleceu dois princípios de transmissão genética: (1) os alelos de um gene segregam um do outro e (2) os alelos de genes diferentes distribuem-se independentemente. O achado de que os genes estão situados em cromossomos possibilitou explicar estes princípios (bem como as exceções a eles) em termos do comportamento meiótico dos cromossomos.

O Princípio de Segregação

(Fig. 5.7). Durante a primeira divisão meiótica, os cromossomos homólogos formam um par. Um dos homólogos veio da mãe e o outro veio do pai. Se a mãe era homozigota para um alelo, A , de um gene nesse cromossomo e o pai era homozigoto para um alelo diferente, a , do mesmo gene, a prole deve ser heterozigota, isto é, Aa . Na anáfase da primeira divisão meiótica, os cromossomos pareados separam-se e se movem para pólos opostos da célula. Um leva o alelo A e o outro o alelo a . Essa separação física dos dois cromossomos segrega os alelos um do outro. Eles, portanto, residirão em células filhas diferentes. O Princípio de Mendel da Segregação é, portanto, baseado na separação dos cromossomos homólogos durante a anáfase da primeira divisão meiótica.

O Princípio da Distribuição Independente

(Fig. 5.8). O Princípio da Distribuição Independente é também baseado nesta separação anafásica. Para compreender a relação, precisamos considerar genes em dois pares diferentes de cromossomos. Suponha que um heterozigoto $Aa Bb$ foi produzido por reprodução de uma fêmea $AA BB$ com um macho $aa bb$. Suponha também que os dois genes estão em cromossomos diferentes. Durante a prófase da meiose I, os cromossomos com alelos A e a formam um par, como também os cromossomos com os alelos B e b . Na metáfase, os dois pares assumem posições no fuso meiótico em preparação para a separação que virá na anáfase. Como existem dois pares de cromossomos, há dois alinhamentos metafásicos distintos:

$$\frac{A}{a} \quad \frac{B}{b} \quad \text{ou} \quad \frac{A}{a} \quad \frac{b}{B}$$

Cada um destes alinhamentos é igualmente possível. Aqui o espaço separa pares diferentes de cromossomos, e a barra separa os membros homólogos de cada par. Durante a anáfase, os alelos acima das barras irão mover-se para um pólo e os alelos abaixo delas irão mover-se para o outro. Quando ocorre disjunção, há portanto 50% de chance de que os ale-

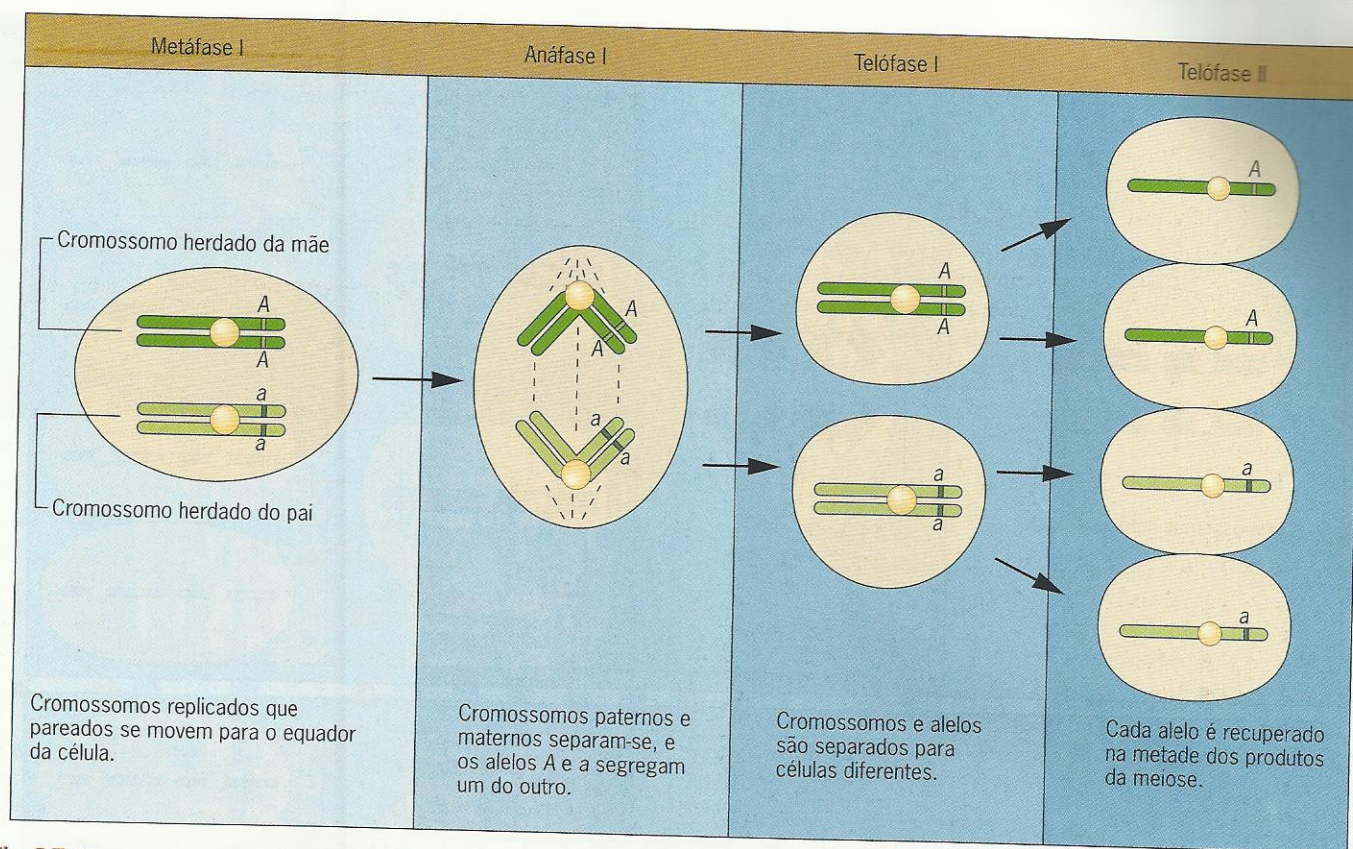


Fig. 5.7 ■ O Princípio da Segregação de Mendel e o comportamento cromossômico meiótico. A segregação dos alelos corresponde à separação de cromossomos pareados na anáfase da primeira divisão meiótica.

los *A* e *B* se movam para o mesmo pólo e 50% de chance de que se movam para pólos opostos. Similarmente, há 50% de chance de que os alelos *a* e *b* se movam para o mesmo pólo e 50% de chance de que se movam para pólos opostos. No final da meiose, quando o número de cromossomos é finalmente reduzido, metade dos gametas deve conter uma combinação parental de alelos (*AB* ou *ab*) e metade deve conter uma nova combinação (*Ab* ou *aB*). Juntos, existirão quatro tipos de gametas, cada um com um quarto do total. Esta igualdade de freqüências gaméticas é um resultado do comportamento independente dos dois pares de cromossomos durante a primeira divisão meiótica. O Princípio da Segregação Independente de Mendel é portanto relativo ao alinhamento aleatório de pares diferentes de cromossomos na metáfase. No Cap. 7, veremos que os genes no mesmo par de cromossomos não segregam independentemente. Em vez disso, como eles estão fisicamente ligados uns aos outros, eles tendem a ser levados juntos na meiose, violando o Princípio da Distribuição Independente.

■ PONTOS IMPORTANTES

- Os genes estão situados em cromossomos.
- A disjunção dos cromossomos durante a meiose é responsável pela segregação e pela distribuição independente dos genes.
- A não-disjunção durante a meiose resulta em números anormais de cromossomos nos gametas e por fim nos zigotos.

GENES LIGADOS AO SEXO EM SERES HUMANOS

Os genes ligados ao X e ao Y foram estudados em humanos.

O desenvolvimento da teoria cromossômica dependeu da descoberta da mutação olho *branco* (*white*) em *Drosophila*. A análise subsequente demonstrou que essa mutação era um alelo recessivo de um gene ligado ao X. Embora alguns de nós possam creditar este episódio importante na história da genética a uma extraordinária boa sorte, a descoberta de Morgan da mutação olho *branco* não foi tão marcante. Tais mutações estão entre as mais fáceis de detectar porque elas se apresentam imediatamente em homens hemizigotos. Em contraste, as mutações autossômicas recessivas apresentam-se apenas após dois alelos mutantes terem sido reunidos em um homozigoto – um evento muito mais improvável.

Em seres humanos, as características recessivas ligadas ao X são muito mais facilmente identificadas do que as autossômicas recessivas. Um homem só precisa herdar um alelo recessivo para mostrar uma característica ligada ao X. Entretanto, uma mulher precisa herdar dois, um de cada um de seus genitores. Assim, a preponderância de pessoas que apresentam características ligadas ao X é de homens.

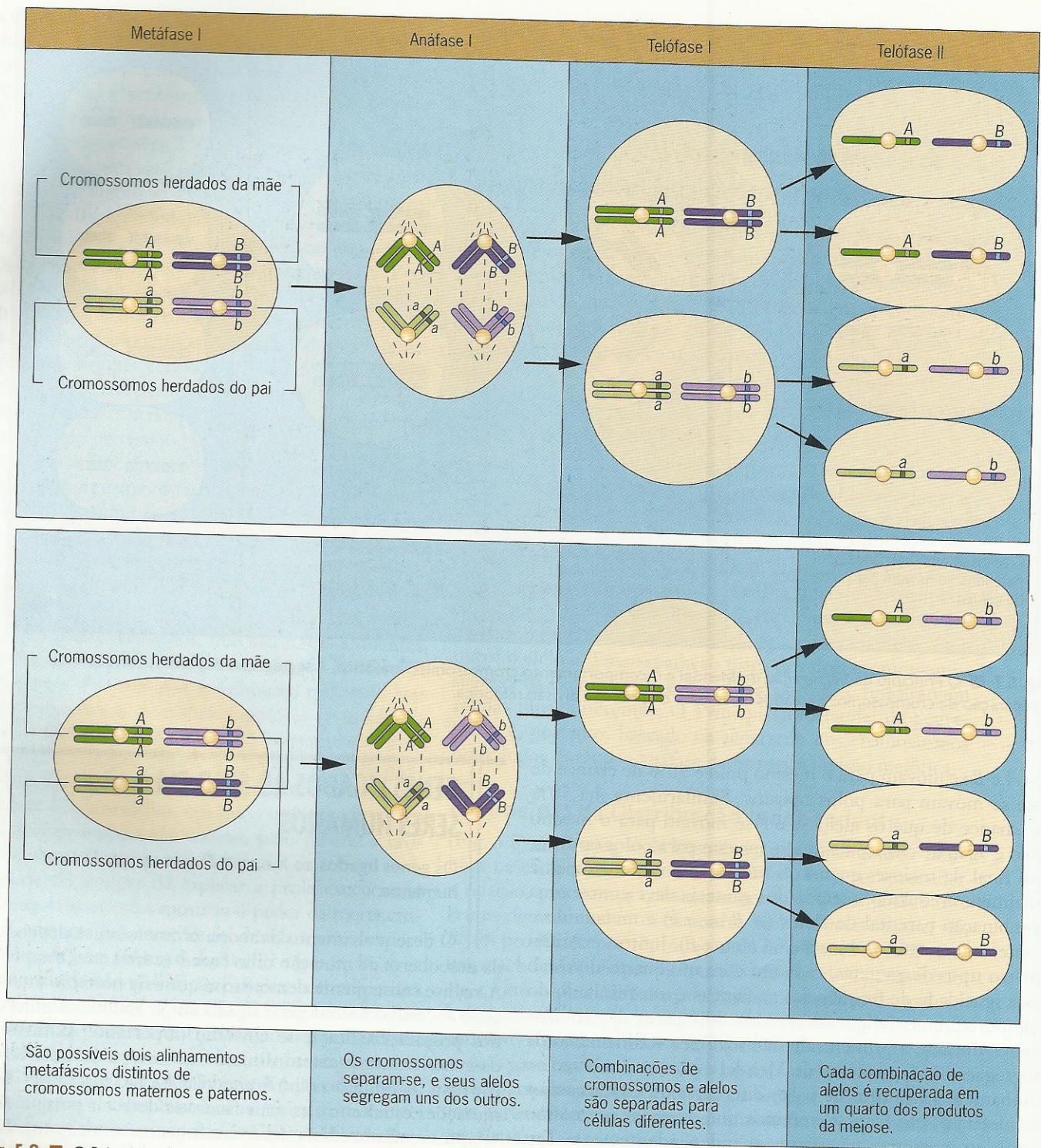


Fig. 5.8 ■ O Princípio da Distribuição Independente de Mendel e o comportamento cromossômico meiótico. Os alelos de pares diferentes de cromossomos distribuem-se independentemente na anáfase da primeira divisão meiótica porque os cromossomos herdados maternamente e paternamente se alinham aleatoriamente no equador da célula.

Hemofilia, um Distúrbio da Coagulação Sangüínea Ligado ao X

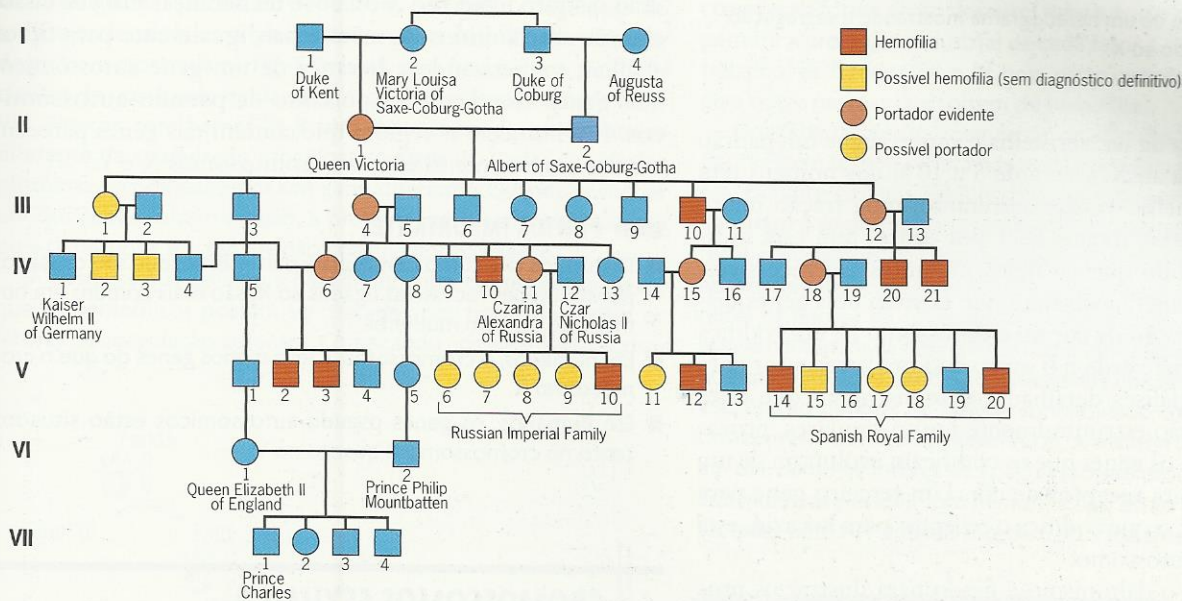
Em seres humanos, um certo tipo de **hemofilia** é um dos melhores exemplos de uma característica ligada ao X. Pes-

soas com esta doença são incapazes de produzir um fator necessário para coagulação sangüínea; cortes, contusões e feridas em hemofílicos continuam a sangrar e, se não forem estancados por tratamento terapêutico, podem causar morte. Quase todas as pessoas afetadas por hemofilia ligada ao X



(a)

Fig. 5.9 ■ (a) A família imperial russa do czar Nicholas II. (b) Hemofilia ligada ao X nas famílias reais da Europa. Através de entrecruzamento, o alelo mutante para hemofilia foi transmitido pela família real britânica para a da Alemanha, da Rússia e da Espanha.



(b)

são homens. Outros distúrbios de coagulação sangüínea são encontrados tanto em homens quanto em mulheres porque são devidos a mutações em genes autossômicos.

O caso mais famoso de hemofilia ligada ao X ocorreu na família imperial russa no começo do século vinte (Fig. 5.9). O czar Nicholas e a czarina Alexandra tiveram quatro filhas e um filho. O filho, Alexis, sofreu de hemofilia. A mutação ligada ao X responsável pela doença de Alexis foi transmitida para ele por sua mãe, que era uma portadora heterozigota. A czarina Alexandra era neta da rainha Victoria do Reino Unido, que também era uma portadora. Os registros de heredogramas mostram que Victoria transmitiu o alelo mutante para três de sua prole de nove: Alice, que era mãe de Alexandra, Beatrice, que teve dois filhos com a doença, e Leopold, que tinha a doença. O alelo de que Victoria era portadora evidentemente surgiu como uma nova mutação em suas células germinativas, ou nas de sua mãe, de seu pai ou de um ancestral materno mais distante.

Ao longo da história, a hemofilia sempre foi uma doença fatal. A maioria das pessoas que a tiveram morreram antes da idade de 20 anos. Hoje em dia, devido à disponibilidade de tratamentos eficazes e relativamente baratos, os hemofílicos têm vida longa e saudável. As terapias que hoje são usadas para tratar hemofilia foram desenvolvidas aplicando-se tecnologias genéticas modernas. O Enfoque Técnico: Hemofilia discute tais avanços médicos.

Daltonismo, um Distúrbio de Visão Ligado ao X

Nos seres humanos, a percepção de cores é mediada por proteínas que absorvem luz nas células cones especializadas da retina do olho. Três destas proteínas foram identificadas, uma que absorve a luz azul, uma que absorve a luz verde e uma que absorve a luz vermelha. O daltonismo pode ser causado por uma anormalidade em qualquer destas proteínas receptoras. O tipo clássico de daltonismo, envolvendo per-

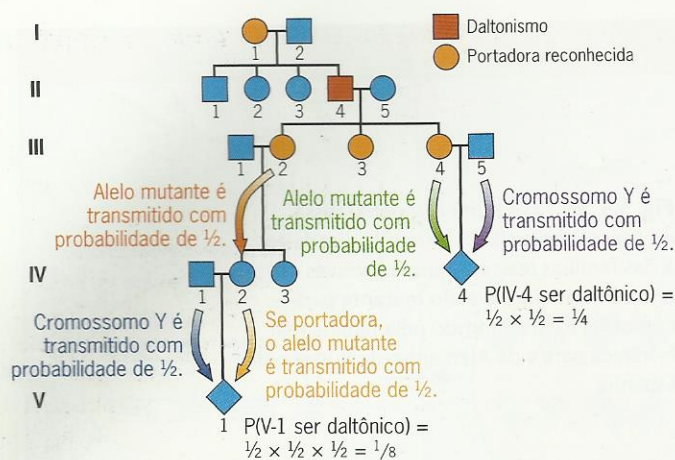


Fig. 5.10 ■ Análise de um heredograma mostrando a segregação do daltonismo ligado ao X.

cepção defeituosa da luz vermelha e verde, segue um padrão de herança ligada ao X. Cerca de 5 a 10% dos homens têm daltonismo vermelho-verde; entretanto, uma fração muito menor de mulheres, menos de 1 por cento, tem esta incapacidade, sugerindo que os alelos mutantes são recessivos. Estudos moleculares mostraram que existem dois genes distintos para percepção de cor no cromossomo X; um codifica o receptor para luz verde e o outro codifica o receptor para luz vermelha. Análises detalhadas demonstraram que estes dois receptores são estruturalmente muito similares, provavelmente porque os genes que os codificam evoluíram de um gene ancestral para receptor de cor. Um terceiro gene para percepção de cor, o que codifica o receptor para luz azul, está situado em um autossomo.

Na Fig. 5.10, o daltonismo é usado para ilustrar os procedimentos destinados a calcular o risco de herdar uma condição recessiva ligada ao X. Um portador heterozigoto, tal como III-4 na figura, tem uma chance de 1/2 de transmitir o alelo mutante para seus filhos. Entretanto, o risco de que um determinado filho seja daltônico é de apenas 1/4, pois a criança teria de ser um menino para manifestar a característica. A mulher indicada como IV-2 no heredograma poderia ser uma portadora do alelo mutante para daltonismo porque sua mãe era. Essa incerteza sobre o genótipo de IV-2 introduz um outro fator de 1/2 no risco de ter um filho daltônico. Assim, o risco para seu filho é de $1/4 \times 1/2 = 1/8$.

Genes no Cromossomo Y Humano

O Projeto do Genoma Humano identificou cerca de 250 genes no cromossomo Y humano. Por comparação, ele identificou mais de 1.000 genes no cromossomo X humano. Antes do trabalho do Projeto do Genoma Humano, pouco se sabia sobre a constituição genética do cromossomo Y. Apenas um punhado de características ligadas ao Y haviam sido detectadas, muito embora a transmissão de pai para filho deva facilitar a identificação na análise convencional de heredogramas.

Os resultados do Projeto do Genoma Humano deram uma explicação possível para a aparente falta de características ligadas ao Y. Vários genes no cromossomo Y humano parecem ser necessários para a fertilidade masculina. Obviamente, uma mutação em tal gene irá interferir na capacidade de um homem de reproduzir. Assim, essa mutação terá pouca ou nenhuma chance de ser transmitida para a geração seguinte.

Genes nos Cromossomos X e Y Humanos

Alguns genes estão presentes tanto nos cromossomos X quanto Y, principalmente perto das pontas dos braços curtos (veja Fig. 5.2). Os alelos desses genes não seguem um padrão distinto ligado ao X ou ao Y de herança. Em vez disso, eles são transmitidos de mães e pais igualmente para filhos e filhas, mimetizando a herança de um gene autossômico. Tais genes são chamados portanto de **pseudo-autossômicos**. Em homens, as regiões que contêm tais genes parecem mediar o pareamento entre os cromossomos X e Y.

PONTOS IMPORTANTES

- Distúrbios tais como hemofilia e daltonismo, que são causados por mutações recessivas ligadas ao X, são mais comuns em homens do que em mulheres.
- Em humanos, o cromossomo Y leva menos genes do que o cromossomo X.
- Em humanos, os genes pseudo-autossômicos estão situados tanto no cromossomo X quanto no Y.

CROMOSSOMOS SEXUAIS E DETERMINAÇÃO DO SEXO

Em alguns organismos, os cromossomos, em particular os cromossomos sexuais, determinam os fenótipos masculino e feminino.

No reino animal, o sexo é talvez o fenótipo mais conspícuo. Os animais com machos e fêmeas distintos são sexualmente dimórficos. Às vezes, esse dimorfismo é estabelecido por fatores ambientais. Em uma espécie de tartarugas, por exemplo, o sexo é determinado pela temperatura. Ovos que foram incubados acima de 30°C desenvolvem-se em fêmeas, enquanto ovos que foram incubados em temperatura mais baixa desenvolvem-se em machos. Em muitas outras espécies, o dimorfismo sexual é estabelecido por fatores genéticos, em geral envolvendo um par de cromossomos sexuais.

Determinação do Sexo em Seres Humanos

A descoberta de que mulheres são XX e homens são XY sugeriu que o sexo deveria ser determinado pelo número de



ENFOQUE TÉCNICO: Hemofilia

Existem dois tipos principais de hemofilia. A forma clássica da doença é a hemofilia A, às vezes chamada “hemofilia real” porque ocorreu nos descendentes da rainha Victoria. A outra forma é a hemofilia B, às vezes chamada “doença de Christmas” porque foi primeiro observada em um paciente chamado de Stephen Christmas. Embora ambos os tipos de hemofilia sejam ligados ao X, eles são causados por mutações em genes diferentes no cromossomo X. A hemofilia A contribui com cerca de 80% de todos os casos de hemofilia ligada ao X.

Ambos os tipos de hemofilia resultam de defeitos na coagulação sanguínea – a cascata de reações que causa coagulação do sangue no local de uma ferida. Uma versão simplificada de parte desta via é mostrada na Fig. 1. Pessoas com hemofilia A são deficientes de uma proteína da coagulação sanguínea chamada fator VIII. Pessoas com hemofilia B têm deficiência de uma proteína diferente da coagulação sanguínea, o fator IX. Cada uma destas proteínas é codificada por um gene diferente. Quando qualquer um dos genes sofre mutação, a proteína codificada está ausente ou é defeituosa, e o sangue não coagula normalmente.

A pesquisa que revelou a base bioquímica da hemofilia sugere que os hemofílicos possam ser tratados por transfusões com os fatores da coagulação ausentes. Começando nos anos 1960, as pro-

teínas de coagulação foram purificadas do sangue obtido de um grande número de doadores e administradas a hemofílicos de forma concentrada. Este processo era caro, e os fatores concentrados em geral não estavam disponíveis para hemofílicos em muitos países. Felizmente, avanços na engenharia genética forneceram meios mais simples, mais baratos e mais confiáveis de obter os fatores cruciais da coagulação. Os genes normais que codificam os fatores VIII e IX foram isolados de amostras de DNA, e cada gene foi introduzido nas células que podiam ser cultivadas no laboratório. Linhagens celulares capazes de sintetizar grandes quantidades de fatores da coagulação foram então desenvolvidas a partir dessas culturas para permitir a produção industrial de cada fator. Como resultado desse trabalho, os fatores VIII e IX hoje estão prontamente disponíveis para tratar pessoas que sofrem de hemofilia.

O uso de fatores da coagulação produzidos nas células cultivadas também tornou o tratamento da hemofilia mais seguro do que era. Durante os anos 1980, muitas pessoas com hemofilia desenvolveram a síndrome da imunodeficiência adquirida, AIDS, porque receberam fatores da coagulação que haviam sido colhidos de doadores de sangue. As doações de sangue eram acumuladas para se obter os fatores da coagulação em quantidade, e parte do sangue estava contaminada com o vírus da imunodeficiência humana (HIV), que causa AIDS. Pacientes com hemofilia que receberam fatores da coagulação obtidos de sangue contaminado tornaram-se infectados com HIV e desenvolveram AIDS. A maioria deles morreu. Hoje em dia, como os fatores da coagulação são produzidos em células cultivadas, as pessoas com hemofilia não estão em risco de ser infectadas pelo HIV quando recebem seus tratamentos.

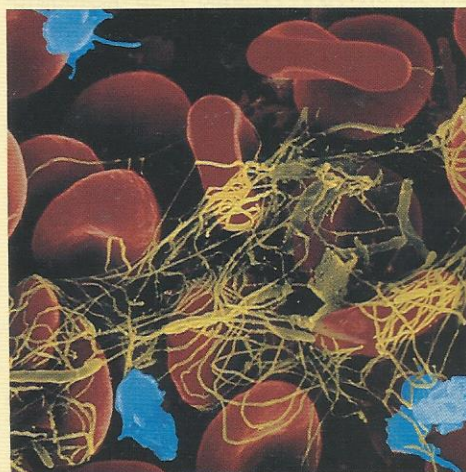
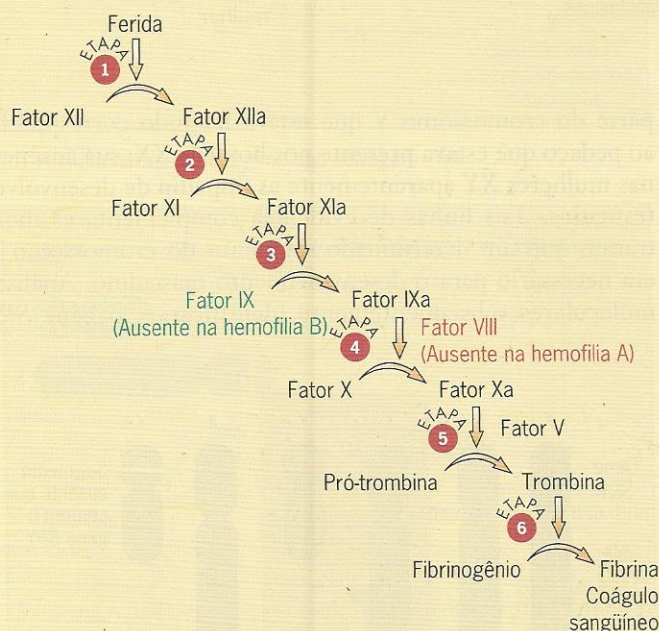


Fig. 1 ■ Via simplificada de coagulação sanguínea em humanos. A cascata é ativada por dano celular em um local ferido. Cada fator é o produto de um gene específico. Em cada etapa da via, um fator é convertido em sua forma ativa (indicada pela letra *a*), que então ativa o fator seguinte na via. Alguns fatores, tais como o fator VIII, participam da ativação de outros fatores sem uma ativação prévia. Nas últimas duas etapas da via, os fatores pró-trombina e fibrinogênio são convertidos em suas formas ativas trombina e fibrina, respectivamente. Longos filamentos de fibrina estabilizam coágulos sanguíneos envolvendo células chamadas plaquetas (foto acima). A ausência de fatores VIII e IX em pessoas com hemofilia A e B, respectivamente, bloqueia o processo de coagulação antes da formação de fibrina, o que resulta em sangramento excessivo ou hemorragia por um corte ou machucado.

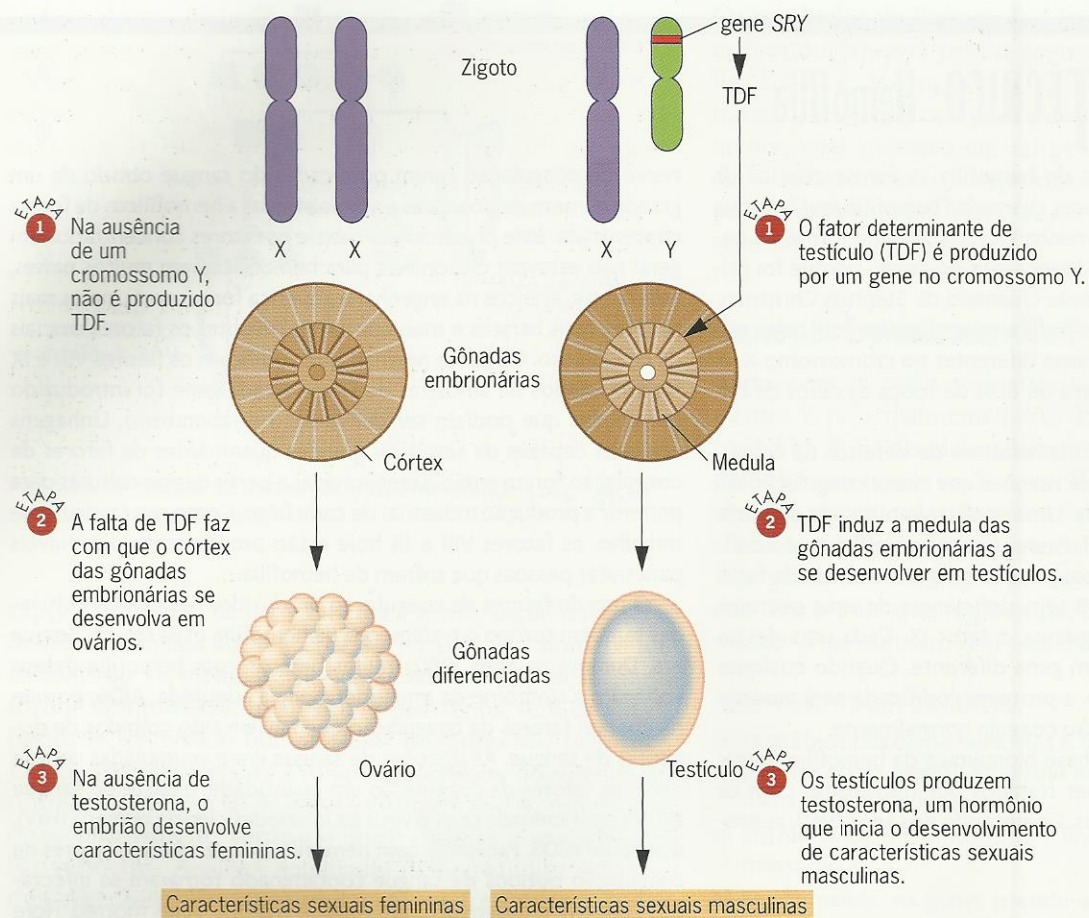


Fig. 5.11 ■ O processo de determinação do sexo em seres humanos. O desenvolvimento sexual masculino depende da produção do fator determinante do sexo (TDF) por um gene no cromossomo Y. Na ausência deste fator, o embrião se desenvolve como mulher.

cromossomos X ou pela presença ou ausência de um cromossomo Y. Como sabemos agora, a segunda hipótese é a correta. Em humanos e outros mamíferos placentários, o sexo masculino é devido a um efeito dominante do cromossomo Y (Fig. 5.11). A evidência deste fato vem do estudo de pessoas com um número anormal de cromossomos sexuais. Animais XO desenvolvem-se em fêmeas e animais XXY desenvolvem-se em machos. O efeito dominante do cromossomo Y manifesta-se cedo no desenvolvimento, quando ele direciona as gônadas primordiais para se desenvolverem em testículos. Uma vez formados os testículos, eles secretam testosterona, um hormônio que estimula o desenvolvimento das características sexuais secundárias masculinas.

Pesquisadores mostraram que o **fator determinante de testículos (TDF)** é o produto de um gene chamado **SRY** (de *sex-determining region Y*), que está situado fora da região pseudo-autossômica no braço curto do cromossomo Y. A descoberta do **SRY** foi possível pela identificação de indivíduos incomuns cujo sexo era inconsistente com sua constituição cromossômica – homens XX e mulheres XY (Fig. 5.12). Foi observado que alguns dos homens XX têm um pequeno trecho do cromossomo Y inserido em um dos cromossomos X. Esse trecho evidentemente tinha um gene responsável pelo desenvolvimento masculino. Foi descoberto que algumas mulheres XY têm um cromossomo Y incompleto. A

parte do cromossomo Y que estava faltando correspondia ao pedaço que estava presente nos homens XX; sua ausência nas mulheres XY aparentemente as impediu de desenvolver testículos. Tais linhas de evidência complementares mostraram que um determinado segmento do cromossomo Y era necessário para o desenvolvimento masculino. Análises moleculares subsequentemente identificaram o gene **SRY**

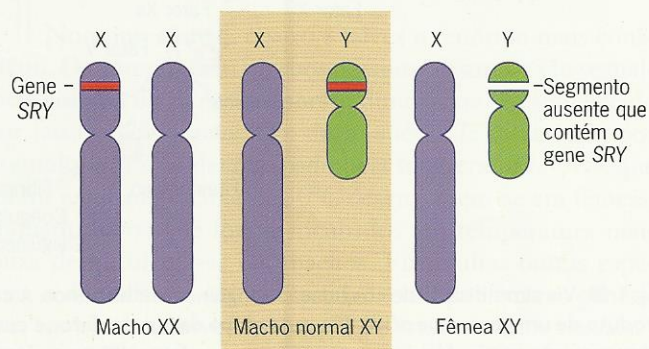
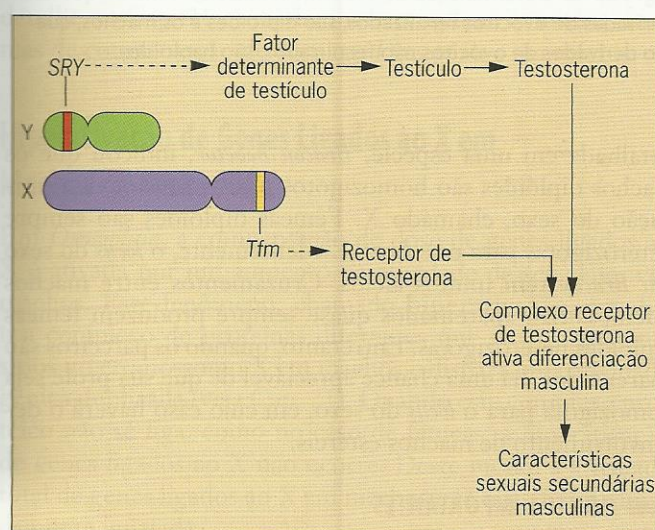


Fig. 5.12 ■ Evidência da localização do gene para o fator determinante de testículo (TDF) no braço curto do cromossomo Y em homens normais. TDF é o produto do gene **SRY**. Em homens XX, uma pequena região contendo este gene foi inserida em um dos cromossomos X e, em mulheres XY, ele foi deletado do cromossomo Y.

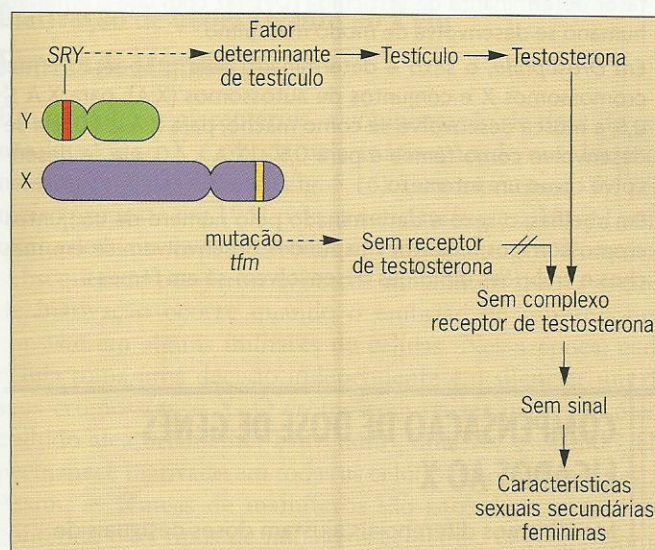
nesse segmento determinante do sexo masculino. Pesquisas adicionais mostraram que um gene *SRY* está presente no cromossomo Y do camundongo e que, como o gene *SRY* humano, ele especifica o desenvolvimento masculino.

Após terem sido formados os testículos, a secreção de testosterona inicia o desenvolvimento das características sexuais masculinas. A testosterona é um hormônio que se liga a receptores em muitos tipos de células. Uma vez ligado, o complexo hormônio-receptor transmite um sinal para o núcleo, instruindo a célula sobre a maneira como se diferenciar. A diferenciação conjunta de muitos tipos de células promove o desenvolvimento de características masculinas distintas tais como musculatura forte, barba e voz grossa. Se o sistema de sinalização de testosterona falha, estas ca-

racterísticas não aparecem e a pessoa se desenvolve como mulher. Um motivo para a falha consiste em incapacidade de produzir o receptor de testosterona (Fig. 5.13). Indivíduos XY com esta deficiência bioquímica inicialmente se desenvolvem de modo masculino – são formados os testículos e é produzida testosterona. Entretanto, a testosterona não exerce efeito porque não pode transmitir o sinal relacionado a desenvolvimento em suas células alvos. As pessoas sem o receptor de testosterona adquirem, portanto, características sexuais femininas. Entretanto, elas não desenvolvem ovários e conseqüentemente são estéreis. Esta síndrome, chamada **feminização testicular**, resulta de uma mutação em um gene ligado ao X, *Tfm*, que codifica o receptor de testosterona. A mutação *tfm* é transmitida pelas mães para sua prole hemizigota XY (que são fenotipicamente fêmeas) em um padrão típico ligado ao X.



(a) Homem normal com o gene tipo selvagem *Tfm*.



(b) Macho com mutação *tfm* e feminização testicular.

Fig. 5.13 ■ Feminização testicular, uma condição causada por uma mutação ligada ao X, *tfm*, que impede a produção do receptor de testosterona.

Determinação do Sexo em *Drosophila*

O cromossomo Y em *Drosophila*, ao contrário do cromossomo Y em humanos, não desempenha papel na determinação do sexo. Em vez disso, o sexo da mosca é determinado pela proporção entre cromossomos X e autossomos. Este mecanismo foi primeiro demonstrado por Bridges em 1921 por meio de uma análise de moscas com constituições cromossômicas incomuns.

Moscas diplóides normais têm um par de cromossomos sexuais, seja XX ou XY, e três pares de autossomos, geralmente representados como AA. Aqui, cada A representa um conjunto haplóide de autossomos. Em experimentos complexos, Bridges obteve moscas com números anormais de cromossomos (Quadro 5.2). Ele observou que, quando a proporção de X para A era de 1,0 ou maior, a mosca era fêmea e, sempre que era de 0,5 ou menor, a mosca era macho. As fêmeas com uma proporção X:A entre 0,5 e 1,0 desenvolviam características de ambos os sexos. Assim, Bridges

QUADRO 5.2

Proporção de Cromossomos X para Autossomos e o Fenótipo Correspondente em <i>Drosophila</i>		
Cromossomos X (X) e Conjuntos de Autossomos (A)	Proporção X:A	Fenótipo
1X 2A	0,5	Masculino
2X 2A	1,0	Feminino
3X 2A	1,5	Metafêmea
4X 3A	1,33	Metafêmea
4X 4A	1,0	♀ tetraplóide
3X 3A	1,0	♀ triplóide
3X 4A	0,75	Intersexo
2X 3A	0,67	Intersexo
2X 4A	0,5	♂ tetraplóide
1X 3A	0,33	Metamacho

as chamou de intersexos. Em nenhuma dessas moscas o cromossomo Y teve qualquer efeito no fenótipo sexual. Ele era, entretanto, necessário para a fertilidade masculina. Os detalhes moleculares da determinação do sexo em *Drosophila* são discutidos nos Caps. 21 e 22.

Determinação do Sexo em Outros Animais

Tanto em *Drosophila* quanto em seres humanos, os machos produzem dois tipos de gametas, portadores de X e portadores de Y. Por este motivo, eles são chamados de sexo **heterogamético**. Nestas espécies, as fêmeas são o sexo **homogamético**. Em aves, borboletas e alguns répteis, tal situação é revertida (Fig. 5.14). Os machos são homogaméticos (geralmente representados por ZZ) e as fêmeas são heterogaméticas (ZW). Entretanto, pouco se sabe sobre o mecanismo de determinação do sexo no sistema de sexo cromossômico Z-W.

Em abelhas, o sexo é determinado pelo animal ser haplóide ou diplóide (Fig. 5.15). Embriões diplóides, que se desenvolvem de zigotos, tornam-se fêmeas; embriões haplóides, que se desenvolveram de ovos não fertilizados, tornam-se machos. Se determinada fêmea se desenvolve ou não em uma forma reprodutiva (a rainha) depende de como ela foi alimentada quando larva. Neste sistema, uma rainha pode controlar a proporção de machos e fêmeas regulando a proporção de ovos não fertilizados que ela põe. Como esse número é pequeno, a maioria da prole é de fêmeas, embora estéreis, que servem como operárias na colméia. Em um sistema haplo-diplo de determinação sexual, os ovos são produzidos por meiose na rainha e os espermatozoides por mitose no macho. Este sistema garante que os zigotos terão o número diplóide de cromossomos e que os não fertilizados terão o número haplóide.

Algumas vespas têm o método de determinação sexual haplo-diplo. Nestas espécies, machos diplóides às vezes são produzidos, mas eles são sempre estéreis. Análise genética

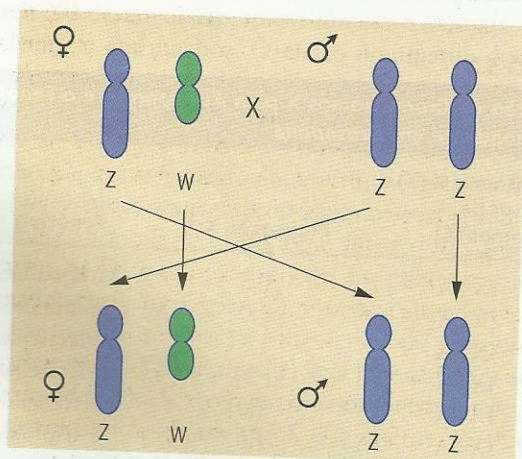


Fig. 5.14 ■ Determinação do sexo em aves. A fêmea é heterogamética (ZW) e o macho é homogamético (ZZ). O sexo da prole é determinado por qualquer um dos cromossomos sexuais, Z ou W, que é transmitido pela fêmea.

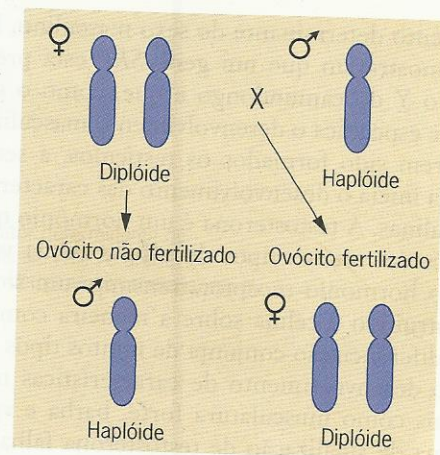


Fig. 5.15 ■ Determinação do sexo em abelhas. As fêmeas, que são derivadas de ovócitos fertilizados, são diplóides e os machos, que são derivados de ovócitos não fertilizados, são haplóides.

detalhada em uma espécie, *Bracon hebetor*, indicou que os machos diplóides são homozigotos para o *locus* de determinação do sexo, chamado X. Fêmeas diplóides são sempre heterozigotas para este *locus*. Evidentemente, o *locus* do sexo em *Bracon* tem muitos alelos. Cruzamentos entre machos e fêmeas não aparentados quase sempre produzem fêmeas diplóides heterozigotas. Entretanto, quando os parceiros são aparentados, há uma chance apreciável de que sua prole seja homozigota para o *locus* do sexo, em cujo caso haverá o desenvolvimento de machos estéreis.

PONTOS IMPORTANTES

- Em humanos, o sexo é determinado por um efeito dominante do gene *SRY* sobre o cromossomo Y; o produto deste gene, o fator determinante de testículo (TDF), faz com que o embrião humano se desenvolva de modo masculino.
- Em *Drosophila*, o sexo é determinado pela proporção entre cromossomos X e conjuntos de autossomos (X:A); para $X:A \leq 0,5$, a mosca desenvolve-se como macho, para $X:A \geq 1,0$, ela se desenvolve como fêmea e para $0,5 < X:A < 1,0$, ela se desenvolve como um intersexo.
- Em abelhas, o sexo é determinado pelo número de conjuntos cromossômicos; embriões haplóides desenvolvem-se em machos e embriões diplóides desenvolvem-se em fêmeas.

COMPENSAÇÃO DE DOSE DE GENES LIGADOS AO X

Mecanismos diferenciais ajustam doses desiguais de genes ligados ao X em animais masculinos e femininos.

O desenvolvimento animal geralmente é sensível a um desequilíbrio no número de genes. Normalmente, cada gene está presente em duas cópias. Desvios desta condição, seja

para cima ou para baixo, podem causar fenótipos anormais e por vezes mesmo a morte. É portanto curioso que tantas espécies devam ter um sistema de determinação do sexo baseado em fêmeas com dois cromossomos X e machos com apenas um. Nestas espécies, de que maneira é acomodada a diferença numérica de genes ligados ao X? *A priori*, três mecanismos podem compensar esta diferença: (1) cada gene ligado ao X poderia funcionar em dobro tanto em machos como nas fêmeas ou (2) uma cópia de cada gene ligado ao X poderia ser inativada em fêmeas ou (3) cada gene ligado ao X poderia funcionar com a metade da intensidade tanto em fêmeas como em machos. Amplas pesquisas mostraram que todos os três mecanismos são usados, o primeiro em *Drosophila*, o segundo em mamíferos e o terceiro no nematódeo *Caenorhabditis elegans*. Tais mecanismos são discutidos em detalhe no Cap. 21. Aqui damos breves descrições dos sistemas de compensação de dose em *Drosophila* e mamíferos.

Hiperativação de Genes Ligados ao X em Machos de *Drosophila*

Em *Drosophila*, a compensação de dose de genes ligados ao X é ativada por um aumento na atividade desses genes em machos. Este fenômeno, chamado **hiperativação**, envolve um complexo de proteínas diferentes que se liga a muitos sítios no cromossomo X em machos e dispara uma atividade gênica duplicada (veja Cap. 21). Quando este complexo proteico não se liga, como no caso de fêmeas, a hiperativação de genes ligados ao X não ocorre. Deste modo, a atividade total de genes ligados ao X em machos e fêmeas é aproximadamente igualada.

Inativação de Genes Ligados ao X em Fêmeas de Mamíferos

Em mamíferos placentários, a compensação de dose de genes ligados ao X é obtida pela **inativação** de um dos cromossomos X das fêmeas (Fig. 5.16). Este mecanismo foi primeiro proposto em 1961 pela geneticista inglesa Mary Lyon, que o deduziu de seus estudos em camundongos. Pesquisas subsequentes por Lyon e outros mostraram que o evento de inativação ocorre quando o embrião de camundongo consiste em alguns milhares de células. Nessa época, cada célula toma uma decisão independente em silenciar um de seus cromossomos X. O cromossomo a ser inativado é escolhido aleatoriamente. Entretanto, uma vez escolhido, ele permanece inativado em todas as células descendentes desta. Assim, as fêmeas de mamíferos são **mosaicos genéticos** contendo dois tipos de linhagens celulares. O cromossomo X herdado maternamente é inativado em cerca de metade destas células, e o X herdado paternamente é inativado na outra metade. Uma fêmea que é heterozigota para um gene ligado ao X é portanto capaz de apresentar dois fenótipos diferentes.

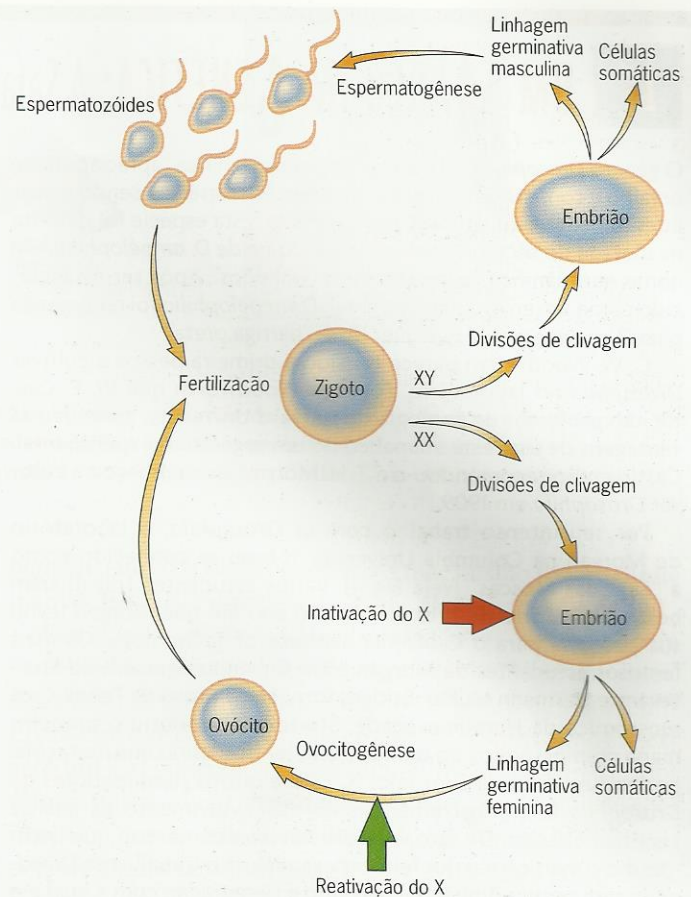


Fig. 5.16 ■ Inativação do cromossomo X em mamíferos. Um dos cromossomos X em fêmeas XX é inativado em cada célula do embrião inicial. Na linhagem germinativa, os cromossomos X inativados são subsequentemente reativados durante a ovocitogênese.

Um dos melhores exemplos deste mosaicismo fenotípico vem do estudo de coloração da pelagem em gatos e camundongos (Fig. 5.17). Em ambas estas espécies, o cromossomo X leva um gene para pigmentação da pelagem. As fêmeas heterozigotas para alelos diferentes deste gene mostram áreas de pelagem clara e escura. As áreas claras expressam um alelo e as áreas escuras expressam o outro. Em gatos, em que um alelo produz pigmento preto e o outro produz pigmento laranja, este fenótipo malhado é chamado de *tortoiseshell*. Cada área de pelagem define um clone de células produtoras de pigmento, ou melanócitos, que foram derivadas por mitose de uma célula precursora presente na época de inativação do cromossomo X.

Um cromossomo X que foi inativado não se parece com outros cromossomos ou age como tais. Análises químicas mostram que seu DNA é modificado pela adição de vários grupos metila. Além disso, ele se condensa em uma estrutura de coloração escura chamada **corpúsculo de Barr** (Fig. 5.18), em homenagem ao geneticista canadense Murray Barr, que primeiro o observou. Esta estrutura torna-se ligada à face interna da membrana nuclear, onde se replica ao final



UM MARCO NA GENÉTICA: A Sala das Moscas de Morgan

O gênero *Drosophila* (da palavra grega que significa “apreciadora de orvalho”) consiste em um grande número de espécies, sendo a mais intensamente estudada a *D. melanogaster*. Esta espécie foi descrita na metade do século dezenove sob o nome de *D. ampelophilia*, um nome que significa “apreciadora de vinhedos”. Após ser estabelecido como um animal experimental, *D. ampelophilia* foi renomeada para *D. melanogaster*, que quer dizer “barriga preta”.

C. W. Woodworth parece ter sido a primeira pessoa a cultivar *Drosophila* no laboratório. Foi com Woodworth que W. E. Castle, um professor de zoologia da Harvard University, aprendeu as vantagens de usar este animal como um organismo experimental. Castle então recomendou-o a T. H. Morgan, que começou a cultivar *Drosophila* em 1909.

Por seu intenso trabalho com as *Drosophila*, o laboratório de Morgan na Columbia University tornou-se conhecido como a “Sala das Moscas” (veja Fig. 1). Vários estudantes trabalharam neste laboratório de 1910 até 1926, o ano em que Morgan levou sua pesquisa para o California Institute of Technology. Os mais famosos estudantes de Morgan eram Calvin Bridges, Alfred Sturtevant e Hermann Muller. Bridges forneceu a prova da Teoria Cromossômica da Hereditariedade, Sturtevant produziu o primeiro mapa cromossômico do mundo e Muller descobriu que mutações podem ser induzidas por raios X. Este e outros pesquisadores de *Drosophila* foram instrumentais no desenvolvimento da análise genética clássica. De fato, William Bateson comentou que “nem mesmo o mais cético dos leitores examinaria o trabalho de *Drosophila* sem sentir admiração pelo zelo e penetração com a qual ele foi conduzido e pela ampla extensão do conhecimento genético ao qual ele levou, mais longe do que qualquer outra linha de trabalho desde os próprios experimentos de Mendel.”¹

Em 1939 Morgan recordou os primeiros dias na Sala das Moscas:

Não era incomum para nós seis trabalhar nesta sala pequena, o único espaço a nossa disposição. Eram dias em que bananas eram usadas como alimento das moscas e em um canto da sala muitas bananas em geral estavam disponíveis, um auxiliar a nossas pesquisas que interessava de outro modo a outros membros do laboratório. Como não havia incubadoras, uma caixa de livros e um cercado tinham lâmpadas e um termostato barato, que às vezes funcionava mal, com conseqüente perda das culturas. O uso de garrafas de leite entrou bem cedo no programa, mas não se sabia de onde vinham, ou pelo menos não foi mencionado... Nossa proximidade uns dos outros levou a uma cooperação em tudo que ocorria.²

Depois, um estudante de Morgan, Alfred Sturtevant, notou que “Havia uma atmosfera de troca na sala das moscas. À medida que surgia cada nova idéia ou resultado, eles eram discutidos livremente pelo grupo.”³ O entusiasmo de usar drosófilas como um organismo experimental e o coleguismo intelectual criado na Sala das Moscas combinavam-se para estimular a pesquisa. O rápido

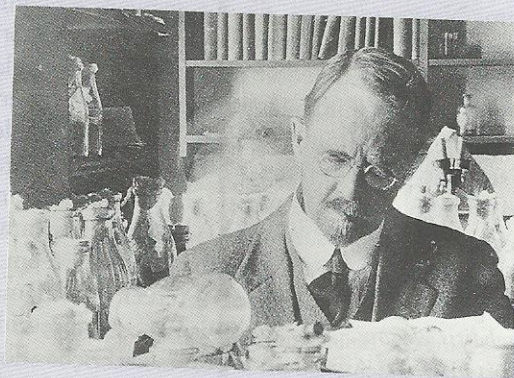


Fig. 1 ■ T. H. Morgan em seu laboratório na Columbia University.

isolamento e análise de novas mutações tornaram a *Drosophila* o principal organismo para a genética experimental.

Desde os dias da Sala das Moscas, as pesquisas com *Drosophila* desenvolveram-se em uma tarefa mundial. Centenas de laboratórios diferentes atualmente estão investigando os genes e os cromossomos deste animal. Nos EUA, um grande centro de estoque mantém culturas de moscas mutantes para distribuição a pesquisadores interessados, e a cada ano pesquisadores de *Drosophila* se juntam num encontro internacional de cinco dias para discutir os resultados de suas pesquisas. Tipicamente, mais de uma centena de pessoas comparecem a esse encontro anual. Em 1992, em um encontro na Philadelphia, foi elaborado um compêndio sobre os 80 anos de pesquisa em *Drosophila*. Esse volume com 1.133 páginas, intitulado *The Genome of Drosophila melanogaster*,⁴ descreve mais de 4.000 genes. As informações sobre os genes de *Drosophila* hoje estão disponíveis em um banco de dados eletrônico regularmente atualizado, chamado FlyBase, que está acessível na Internet pela World Wide Web (<http://www.flybase.org/>). Em 2000, no encontro em Pittsburgh, foi obtido o término do projeto sobre a sequência do DNA para a maioria do genoma de *Drosophila*; 120 milhões de pares de bases constituindo 13.601 genes estimados foram analisados, um legado que certamente faria feliz T. H. Morgan.

QUESTÕES PARA DISCUSSÕES

1. *Drosophila melanogaster* foi a primeira espécie a tornar-se estabelecida como organismo “modelo” para pesquisa genética. Outros organismos de reinos biológicos diferentes foram então elevados a este status especial. Quais as vantagens e desvantagens de concentrar a pesquisa em organismos “modelos”?
2. A Sala das Moscas de Morgan demonstrou o poder do trabalho em equipe em pesquisas científicas e estabeleceu um padrão para a organização laboratorial que muitos cientistas seguiram – estudantes trabalhando com um “pesquisador principal” para estudar aspectos diferentes de um problema de pesquisa. Quais os benefícios e as desvantagens em fazer uma pesquisa deste modo? Quais são as responsabilidades de um pesquisador principal e de seus estudantes em uma equipe de pesquisa?

¹Bateson, W. 1916. The mechanism of Mendelian heredity (a review). *Science* 44:536-543.

²Morgan, T. H. 1939. “Personal recollections of Calvin B. Bridges.” *Journal of Heredity* 30:355.

³Sturtevant, A. H. 1965. *A History of Genetics*. Harper and Row, New York, p. 49.

⁴Lindsley, D. L. e G. G. Zimm. 1992. *The Genome of Drosophila melanogaster*. Academic Press, New York.



Fig. 5.17 ■ Mosaicos de cor resultantes de inativação do cromossomo X em fêmeas de mamíferos. (a) Formação de clones de células em um embrião de gato que produzem manchas diferentes na pelagem do adulto. (b) Um gato *tortoiseshell*. Esta fêmea é heterozigota para um gene de cor de pelagem ligado ao X. As manchas laranja e pretas são devidas a inativação de alelos diferentes nas células produtoras de pigmento do corpo.

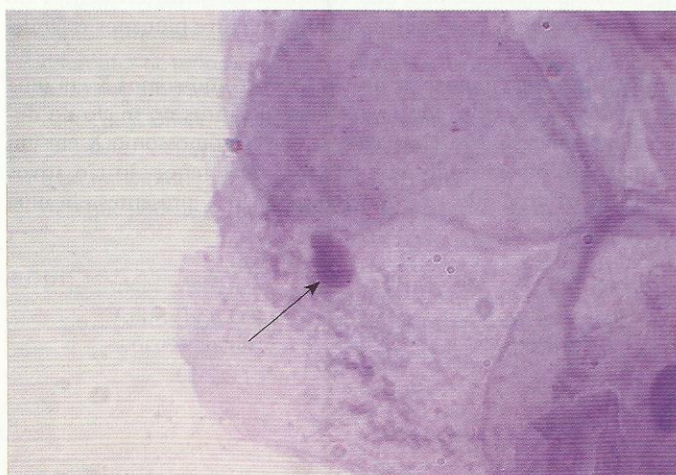


Fig. 5.18 ■ Corpúsculo de Barr em uma célula humana feminina corada com cresil violeta.

do período de replicação dos demais cromossomos da célula. O cromossomo X inativado permanece neste estado alterado em todos os tecidos somáticos. Entretanto, nos tecidos germinativos ele é reativado, talvez porque são necessárias duas cópias de alguns genes ligados ao X para o término

bem-sucedido da ovocitogênese. O mecanismo molecular de inativação do X é discutido no Cap. 21.

Estudos citológicos identificaram seres humanos com mais de dois cromossomos X (veja Cap. 6). Na maior parte, essas pessoas são mulheres normais do ponto de vista fenotípico, aparentemente porque todos os cromossomos X exceto um são inativados. Geralmente, todos os X inativados transformam-se em corpúsculos de Barr. Estas observações sugerem que as células podem ter uma quantidade limitada de algum fator necessário para evitar inativação do X. Uma vez que este fator tenha sido usado para manter um cromossomo X ativo, todos os outros sofrem o processo de inativação.*

PONTOS IMPORTANTES

- Em *Drosophila*, a compensação de dose para genes ligados ao X é obtida por hiperativação do único cromossomo X em machos.
- Em mamíferos, a compensação de dose para genes ligados ao X é obtida por inativação de um dos dois cromossomos X nas fêmeas.

*N.T.: L. Russell demonstrou que a inativação do X não atinge todos os genes, ou seja, não é total.

Exercícios Básicos

Ilustram a análise genética básica.

1. Um macho mutante de *Drosophila* com olhos cor-de-ameixa foi cruzado com uma fêmea tipo selvagem com olhos vermelhos. Toda a prole de F_1 de ambos os sexos tem olhos vermelhos. Quando essa prole foi entrecruzada, foram produzidas três classes diferentes de moscas da F_2 : fêmeas com olhos vermelhos, machos com olhos vermelhos e machos com olhos cor-de-ameixa. Os machos e fêmeas eram igualmente frequentes na F_2 , e, entre os machos, as duas cores de olhos eram igualmente

frequentes. Estes resultados sugerem que a mutação *prune* está no cromossomo X?

Resposta: Os resultados destes cruzamentos são consistentes com a hipótese de que a mutação *prune* está no cromossomo X. De acordo com esta hipótese, o macho do primeiro cruzamento deve ter sido hemizigoto para a mutação *prune*. Esse cruzamento deve ter sido homozigoto para o alelo tipo selvagem do gene *prune*. Entre a F_1 , as filhas devem ter sido heterozigotas para

Variação no Número e na Estrutura dos Cromossomos

PANORAMA

- ▶ Técnicas citológicas
- ▶ Poliploidia
- ▶ Aneuploidia
- ▶ Rearranjos da estrutura do cromossomo

Cromossomos, agricultura e civilização

O cultivo de trigo surgiu há cerca de 10.000 anos no Oriente Médio. Hoje, o trigo é o principal produto agrícola para mais de um bilhão de pessoas. É cultivado em diversos ambientes, desde a Noruega até a Argentina. Desenvolveram-se mais de 17.000 variedades, cada uma delas adaptada a uma região diferente. A produção mundial de trigo é de 60 milhões de toneladas anuais, o que representa mais de 20% das calorias consumidas por toda a população humana. Sem dúvida, o trigo é um produto agrícola importante e, segundo alguns, um sustentáculo da civilização.

O trigo cultivado atualmente, *Triticum aestivum*, é um híbrido de no mínimo três espécies diferentes. Originou-se de gramíneas de baixo rendimento que cresciam na Síria, no Irã, no Iraque e na Turquia. Aparentemente, algumas dessas gramíneas eram cultivadas por povos antigos dessa região. Embora não conheçamos o curso exato dos acontecimentos, parece ter havido um intercruzamento de duas dessas gramíneas, produzindo uma espécie que se destacou como planta de cultura. Graças ao cultivo humano, essa espécie híbrida passou por melhoramento seletivo e também foi intercruzada com uma terceira espécie, produzindo um híbrido triplo, ainda mais adequado para a agricultura. O trigo moderno é descendente dessas plantas híbridas triplas.

O que tornou os trigos triplos-híbridos tão superiores a seus ancestrais? Eles tinham grãos maiores, cresciam em condições mais variadas e a colheita era mais fácil. Agora compreendemos a base cromossômica desses aperfeiçoamentos. O trigo triplo-híbrido contém os cromossomos de cada progenitor. Geneticamente, é uma fusão dos genomas de três espécies diferentes.



Campo de trigo.

Técnicas citológicas

Os geneticistas usam corantes para identificar cromossomos específicos e analisar suas estruturas.

Os geneticistas estudam o número e a estrutura dos cromossomos por coloração das células em divisão com determinados corantes, seguida por exame microscópico. A análise de cromossomos corados é a principal atividade da especialidade chamada **citogenética**.

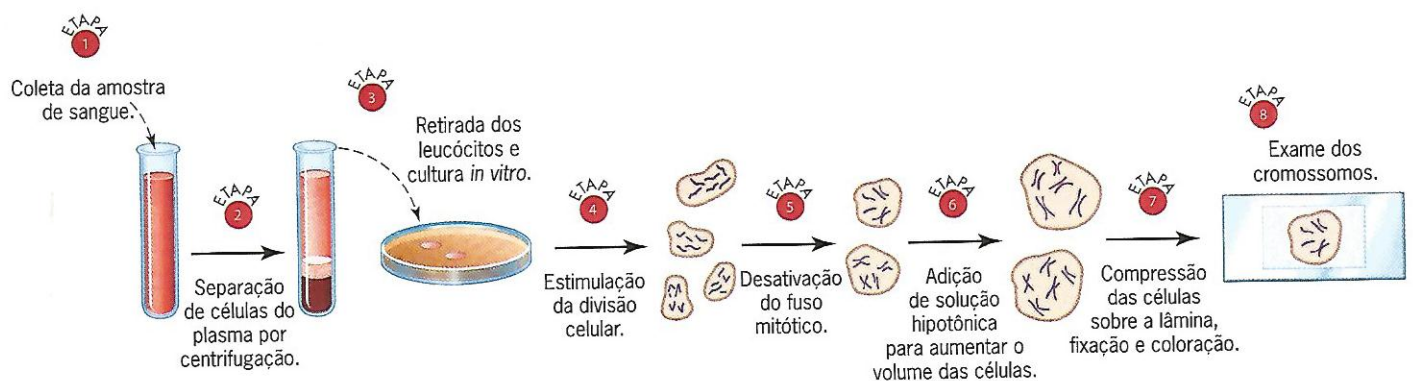
A citogenética originou-se da pesquisa de vários biólogos europeus do século 20 que descobriram os cromossomos e observaram seu comportamento durante a mitose, a meiose e a fertilização. Essa pesquisa prosperou durante o século 20, com o surgimento de microscópios aperfeiçoados e de melhores procedimentos de preparo e coloração dos cromossomos. A demonstração de que os genes estão localizados nos cromossomos fomentou o interesse nessa pesquisa e levou a importantes estudos sobre o número e a estrutura dos cromossomos. Hoje, há aplicação de importantes conhecimentos citogenéticos, principalmente na medicina, em que são usados para identificar a associação entre doenças e anormalidades cromossômicas.

ANÁLISE DE CROMOSSOMOS MITÓTICOS

Os pesquisadores empregam células em divisão, geralmente no meio da mitose, na maioria das análises citológicas. Para aumentar o número de células nesse estágio, costumavam usar material em crescimento, como embriões de animais e extremidades das raízes dos vegetais. O desenvolvimento de técnicas de cultura celular, porém, tornou possível estudar cromossomos em outros tipos de células (**Figura 6.1**). Por exemplo, leucócitos humanos podem ser coletados do sangue periférico, separados das hemácias – que não se dividem – e cultivados. Em seguida, estimula-se a divisão dos leucócitos por tratamento químico e, no meio da divisão, prepara-se uma amostra das células para análise citológica. O procedimento habitual é tratar as células em divisão com

uma substância química que desative o fuso mitótico. Essa interferência captura os cromossomos em mitose, quando é mais fácil observá-los. As células cuja mitose foi interrompida são imersas em solução hipotônica, o que faz com que absorvam água por osmose e aumentem de volume. O conteúdo de cada célula é diluído pela água que entra, de maneira que os cromossomos se dispersam livremente quando elas são comprimidas sobre uma lâmina de microscópio. Essa técnica facilita muito a análise subsequente, sobretudo se houver grande número de cromossomos. Durante muitos anos acreditou-se erroneamente que as células humanas continham 48 cromossomos. O número correto, 46, só foi determinado depois do uso da técnica de aumento do volume celular para separar os cromossomos dentro de cada célula mitótica.

Até o fim da década de 1960 e início da década de 1970, as dispersões cromossômicas geralmente eram coradas com reagente de Feulgen, corante roxo que reage com as moléculas de açúcar no DNA, ou com acetocarmim, um corante vermelho-escuro. Como esses tipos de corantes tingem uniformemente os cromossomos, é impossível para o pesquisador distinguir um cromossomo do outro, a menos que os cromossomos sejam muito diferentes em tamanho ou nas posições de seus centrômeros. Atualmente os citogeneticistas usam corantes que fazem o tingimento diferencial dos cromossomos ao longo de seus comprimentos. A *quinacrina*, substância química semelhante ao antimalárico quinina, foi um dos primeiros reagentes com maior capacidade de discriminação. Os cromossomos corados com quinacrina têm um padrão característico de faixas brilhantes sobre fundo escuro. No entanto, como a quinacrina é uma substância fluorescente, as faixas só aparecem quando os cromossomos são expostos à luz ultravioleta (UV). A irradiação ultravioleta faz com que algumas das moléculas de quinacrina inseridas no cromossomo emitam energia. Partes do cromossomo emitem brilho intenso, enquanto outras continuam escuras. Esse padrão de faixas brilhantes e escuras é altamente reproduzível e também é específico de cada cromossomo (**Figura 6.2**). Desse modo, com o bandeamento com quinacrina, os citogeneticistas são capazes de identificar cromossomos específicos em uma célula e também de identificar



■ FIGURA 6.1 Preparo de células para análise citológica.

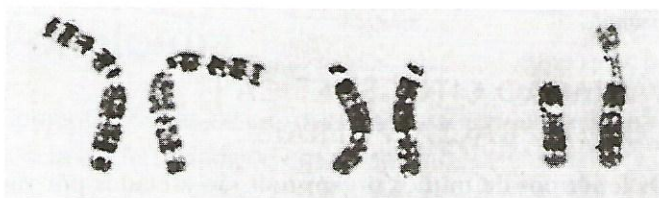


■ FIGURA 6.2 Cromossomos metafásicos da planta *Allium carinatum*, corados com quinacrina.

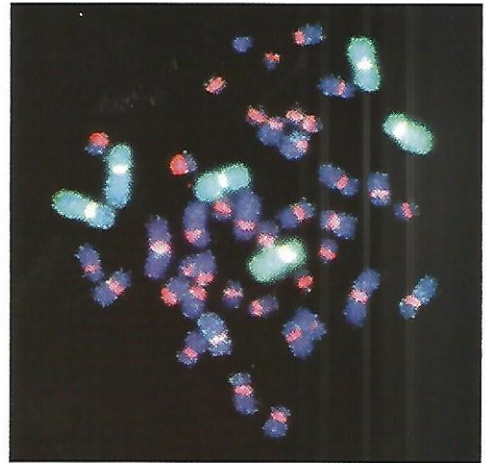
anormalidades na estrutura de um cromossomo, por exemplo, a ausência de algumas bandas.

Também foram desenvolvidas excelentes técnicas de coloração não fluorescente. A mais popular usa o *Giemsa*, mistura de corantes que recebeu esse nome em homenagem ao seu inventor, Gustav Giemsa. Assim como a quinacrina, o Giemsa cria um padrão reproduzível de faixas em cada cromossomo (Figura 6.3). Ainda não está claro por que os cromossomos apresentam faixas quando são corados com quinacrina ou Giemsa. É possível que esses tipos de corantes reajam preferencialmente com determinadas sequências de DNA ou com as proteínas associadas a elas, e que essas sequências de DNA específicas sejam distribuídas de maneira característica dentro de cada cromossomo.

A técnica mais avançada usada por citogeneticistas atualmente é a **pintura cromossômica**. Essa técnica cria imagens coloridas dos cromossomos pelo tratamento das dispersões cromossômicas com fragmentos de DNA, isolados e caracterizados em laboratório, marcados com corante fluorescente. Esse fragmento pode, por exemplo, pertencer a um gene específico. O fragmento do DNA é marcado quimicamente com um corante fluorescente no laboratório e, depois, aplicado aos cromossomos dispersos sobre uma lâmina de vidro. Em condições adequadas, o fragmento de DNA se liga ao DNA cromossômico cuja sequência é complementar à dele. Essa ligação, na verdade, marca o DNA cromossômico com o corante fluorescente presente no fragmento de DNA. Em vista da natureza específica da interação entre o fragmento de DNA e o DNA com-



■ FIGURA 6.3 Cromossomos metafásicos do muntíaco asiático, animal semelhante a um veado, corados com Giemsa.

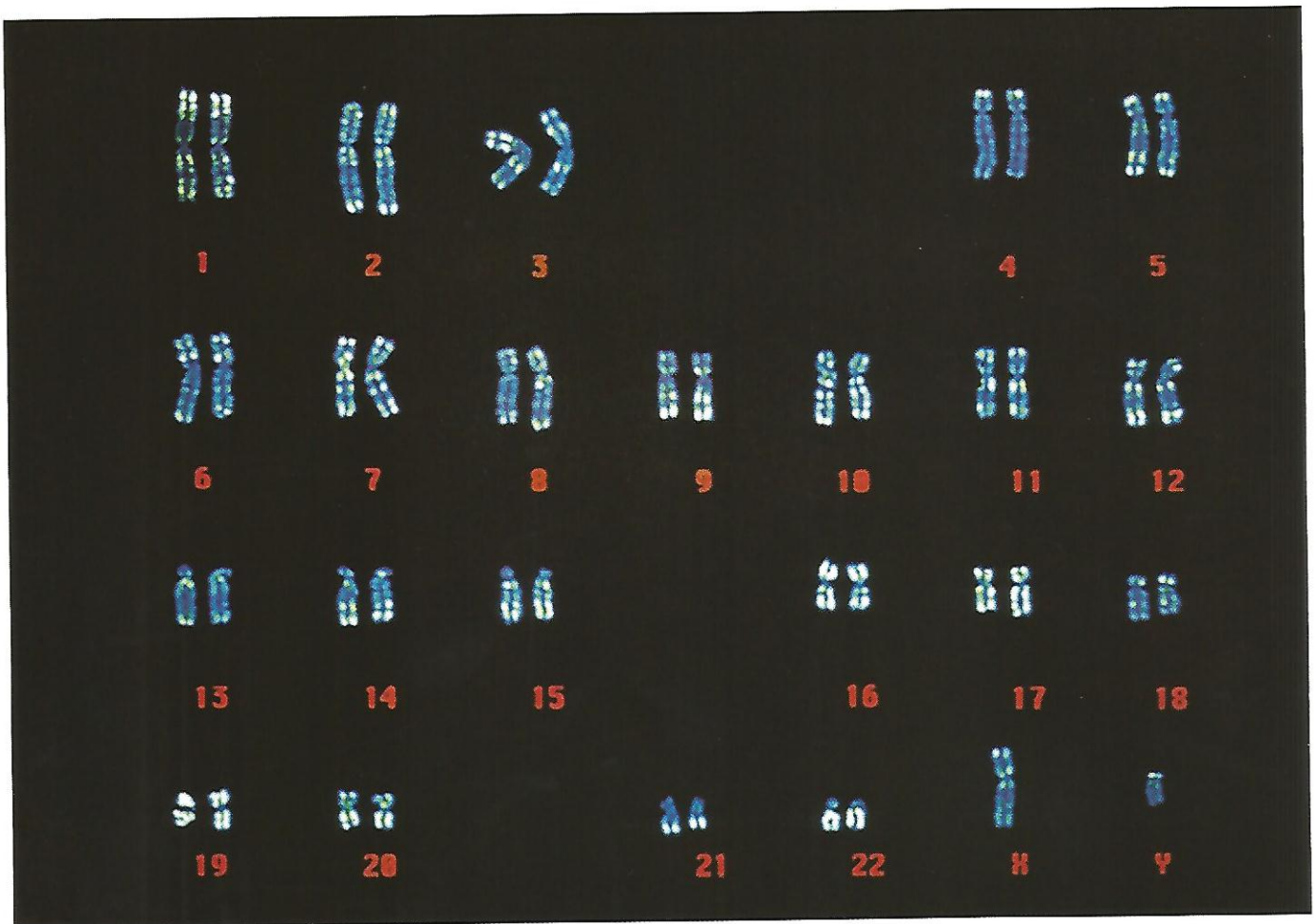


■ FIGURA 6.4 Pintura de cromossomos. Sondas de DNA humano foram aplicadas a uma dispersão de cromossomos humanos. Cada sonda foi marcada com um corante fluorescente de cor diferente (rosa ou verde brilhante) para mostrar a localização cromossômica das sequências de DNA complementares a essas sondas. O alvo da sonda rosa é o DNA no centrômero de todos os cromossomos, enquanto o alvo da sonda verde brilhante é o DNA de apenas três pares de cromossomos.

plementar nos cromossomos, geralmente chamamos o fragmento de DNA de *sonda*. Ele busca seu complemento na grande massa de DNA cromossômico de uma célula e se liga a ele. Depois da ligação da sonda, as dispersões cromossômicas são irradiadas com luz de comprimento de onda apropriado. As faixas ou pontos de cor observados revelam onde está localizada a sequência de DNA complementar – o alvo da sonda – nos cromossomos. A Figura 6.4 mostra cromossomos humanos analisados por essa técnica. Os cromossomos foram pintados simultaneamente com dois fragmentos diferentes de DNA humano, cada um deles marcado com um corante fluorescente de cor diferente. Um dos fragmentos liga-se de maneira inespecífica aos centrômeros de cada cromossomo e, quando estimulado, adquire cor rosa. O outro fragmento liga-se apenas a alguns cromossomos e, quando estimulado, adquire cor verde brilhante. Dessa maneira, esses poucos cromossomos destacam-se entre todos os cromossomos na dispersão. A Figura 2.7 mostra cromossomos humanos pintados com uma série de sondas constituídas de fragmentos de DNA humano. Cada par de cromossomos tem um padrão característico de bandas. Assim, essa técnica torna possível identificar cada par.

CARIÓTIPO HUMANO

As células humanas diploides contêm 46 cromossomos – 44 autossomos e dois cromossomos sexuais, que são XX no sexo feminino e XY no sexo masculino. Na metáfase mitótica, todos os 46 cromossomos são constituídos de duas cromátides-irmãs idênticas. Quando corados apropriadamente, cada cromossomo duplicado pode ser reconhecido pelo tamanho, formato e padrão de bandas. Para análise citológica, as dispersões meta-



■ FIGURA 6.5 Cariótipo corado de um homem para mostrar as bandas de cada cromossomo. Os autossomos são numerados de 1 a 22. X e Y são os cromossomos sexuais.

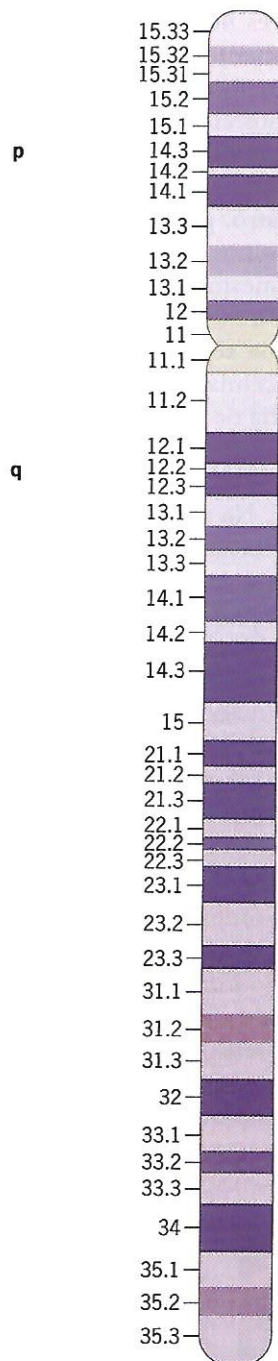
fásicas bem coradas são fotografadas e, depois, a imagem de cada cromossomo é recortada, combinada com seu parceiro para formar pares homólogos, que são organizados em ordem decrescente de tamanho em um quadro (Figura 6.5). O maior autossomo é o número 1, e o menor é o número 21. (Por motivos históricos, o segundo menor cromossomo foi designado número 22.) O cromossomo X tem tamanho intermediário, e o cromossomo Y tem aproximadamente o mesmo tamanho do cromossomo 22. Esse quadro de cromossomos é denominado **cariótipo** (termo originado do grego que significa “núcleo”, em referência ao conteúdo do núcleo). Um pesquisador experiente pode usar o cariótipo para identificar anormalidades do número e da estrutura dos cromossomos.

Antes das técnicas de bandeamento e pintura, era difícil distinguir um cromossomo humano de outro. Os citogeneticistas só poderiam organizar os cromossomos em grupos de acordo com o tamanho, classificando o maior como grupo A, o segundo maior como grupo B, e assim por diante. Embora tenham reconhecido sete grupos diferentes, era quase impossível identificar um cromossomo nesses grupos. Hoje, graças às técnicas de bandeamento e pintura, é possível fazer a identificação rotineira de cada cromossomo. As técnicas de bandeamento e pin-

tura também tornaram possível distinguir cada braço de um cromossomo e investigar regiões específicas neles. O centrômero divide cada cromossomo em braços longo e curto. O braço curto é designado pela letra *p* (do francês *petite*, que significa “pequeno”) e o braço longo, pela letra *q* (porque sucede o “p” no alfabeto). Assim, por exemplo, um citogeneticista pode se referir especificamente ao braço curto do cromossomo escrevendo apenas “5p”. Em cada braço, regiões específicas são designadas por números, a partir do centrômero (Figura 6.6). Assim, no braço curto do cromossomo 5, temos a região 5p11, mais próxima do centrômero, seguida pelas regiões 5p12, 5p13, 5p14 e 5p15, que é a mais distante do centrômero. Dentro de cada região, cada banda é designada por números depois do ponto; por exemplo, 13.1, 13.2 e 13.3 referem-se às três bandas que constituem a região 5p13. O padrão de bandas no cromossomo é denominado *idiotograma*.

VARIAÇÃO CITOGENÉTICA | CONSIDERAÇÕES GERAIS

Os fenótipos de muitos organismos são afetados por variações no número de cromossomos em suas células; às vezes, até mesmo alterações em parte de um cromossomo



■ **FIGURA 6.6** O idiograma do cromossomo 5 humano. As regiões de cada braço são numeradas consecutivamente a partir do centrômero. As bandas de cada região são designadas por números depois de um ponto.

podem ser significativas. Essas diferenças numéricas geralmente são descritas como variações da *ploidia* do organismo (termo de origem grega que significa “vez”, como em “duas vezes”). Organismos com conjuntos completos, ou normais, de cromossomos são euploides (do grego, “bom” e “vez”). Organismos que têm conjuntos adicionais de cromossomos são *poliploides* (do grego, “muitas” e “vezes”), e o nível de poliploidia é descrito referindo-se a um número básico de cromossomos, geralmente designado por n . Assim, diploides, com dois conjuntos básicos de cromossomos, têm $2n$ cromossomos; os triploides, com três conjuntos, têm $3n$; os tetraploides, com quatro conjuntos, têm $4n$; e assim por diante. Os organismos nos quais há deficiência ou excesso de determinado cromossomo, ou segmento de cromossomo, são aneuploides (do grego, “não”, “bom” e “vez”). Portanto, esses organismos sofrem de um desequilíbrio genético específico. Quanto à distinção entre aneuploidia e poliploidia, a aneuploidia é uma alteração numérica em parte do genoma, geralmente em um único cromossomo, enquanto a poliploidia é uma alteração numérica em um conjunto completo de cromossomos. A aneuploidia implica desequilíbrio genético, mas a poliploidia, não.

Os citogeneticistas também catalogaram vários tipos de alterações estruturais nos cromossomos dos organismos. Por exemplo, um fragmento de um cromossomo pode ser fundido a outro cromossomo, ou um segmento dentro de um cromossomo pode ser invertido em relação ao restante desse cromossomo. Essas alterações estruturais são denominadas *rearranjos*. A segregação irregular de alguns rearranjos durante a meiose torna possível associá-los à aneuploidia. Nas seções adiante, abordamos todas essas variações citogenéticas – poliploidia, aneuploidia e rearranjos cromossômicos.

PONTOS ESSENCIAIS

- A análise citogenética geralmente tem como objeto os cromossomos das células em divisão
- Corantes como a quinacrina e Giemsa criam padrões de bandas úteis na identificação individual dos cromossomos em uma célula
- O cariótipo apresenta os cromossomos duplicados de uma célula organizados para análise citogenética.

Poliploidia

Conjuntos extras de cromossomos podem afetar a aparência e a fertilidade de um organismo.

A **poliploidia**, presença de conjuntos extras de cromossomos, é bastante comum em vegetais, mas muito rara

em animais. Metade dos gêneros conhecidos de vegetais contém espécies poliploides, e cerca de dois terços das gramíneas são poliploides. Muitas dessas espécies têm reprodução assexuada. Em animais, nos quais a reprodução é basicamente sexuada, a poliploidia é rara,

provavelmente porque interfere no mecanismo de determinação sexual.

Um efeito geral da poliploidia é o aumento do tamanho da célula, provavelmente porque existem mais cromossomos no núcleo. Com frequência, o aumento de tamanho está relacionado com o aumento geral de tamanho do organismo. As espécies poliploides tendem a ser maiores e mais robustas que as diploides correspondentes. Essas características têm um significado prático para os seres humanos, cuja alimentação depende de muitas espécies vegetais poliploides. Tais espécies tendem a produzir sementes e frutos maiores, portanto, têm maior rendimento agrícola. Trigo, café, batata, banana, morango e algodão são vegetais poliploides, assim como muitas plantas ornamentais cultivadas, entre elas rosas, crisântemos e tulipas (Figura 6.7).

POLIPLOIDES ESTÉREIS

Apesar da aparência física robusta, muitas espécies poliploides são estéreis. Conjuntos extras de cromossomos têm segregação irregular na meiose, com a produção de gametas muito desequilibrados (*i. e.*, aneuploides). Caso haja união desses gametas na fertilização, os zigotos produzidos quase sempre morrem. A inviabilidade dos zigotos explica por que muitas espécies poliploides são estéreis.

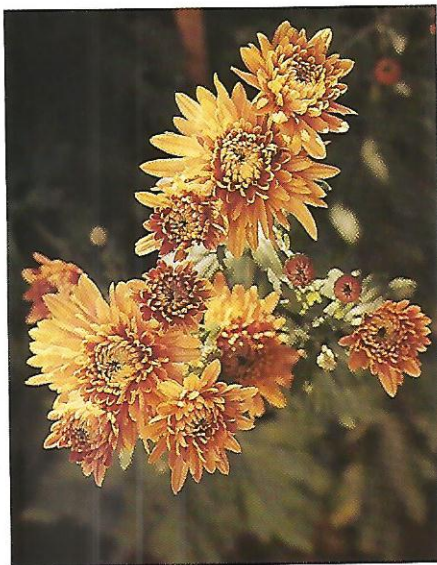
Como exemplo, vamos considerar uma espécie triploide com três conjuntos idênticos de n cromossomos. Portanto, o número total de cromossomos é $3n$. Durante a meiose, cada cromossomo tenta formar par com seu homólogo (Figura 6.8). Uma possibilidade é que haja pareamento de dois homólogos ao longo de todo o comprimento, deixando o terceiro sem par; esse cromossomo solitário é denominado **univalente**. Outra possibilidade

é a sinapse dos três homólogos, formando um **trivalente** no qual há pareamento parcial de cada membro com os dois outros. Em qualquer caso, é difícil prever como será o deslocamento dos cromossomos durante a anáfase da primeira divisão meiótica. O acontecimento mais provável é que dois homólogos sigam para um polo e um homólogo para o outro, produzindo gametas com uma ou duas cópias do cromossomo. Os três homólogos, porém, podem ir para o mesmo polo, produzindo gametas com nenhuma ou com três cópias do cromossomo. Como essa incerteza acerca da segregação aplica-se a cada trio de cromossomos na célula, o número total de cromossomos de um gameta varia de zero a $3n$.

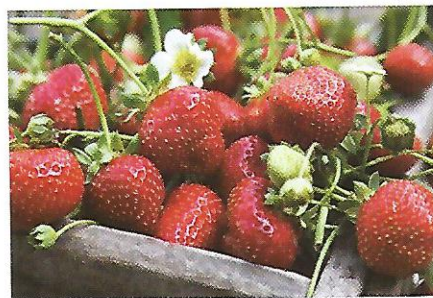
É quase certa a morte dos zigotos formados por fertilização desses gametas; assim, a maioria dos triploides é totalmente estéril. Na agricultura e na horticultura, essa esterilidade é contornada pela propagação assexuada das espécies. Os muitos métodos de propagação assexuada incluem cultivo a partir de estacas (bananas), enxertos (maças Winesap, Gravenstein e Baldwin) e bulbos (tulipas). Na natureza, as plantas poliploides também podem se reproduzir de maneira assexuada. Um mecanismo é a **apomixia**, no qual há meiose modificada com produção de oosferas não reduzidas; essas oosferas formam sementes que germinam e dão origem a novas plantas. O dente-de-leão, uma planta poliploide muito produtiva, reproduz-se dessa maneira.

POLIPLOIDES FÉRTEIS

As incertezas meióticas que ocorrem em triploides também ocorrem em tetraploides, que têm quatro conjuntos idênticos de cromossomos e, portanto, também são estéreis. Alguns tetraploides, porém, são capazes de gerar prole viável. O exame atento mostra que essas espécies



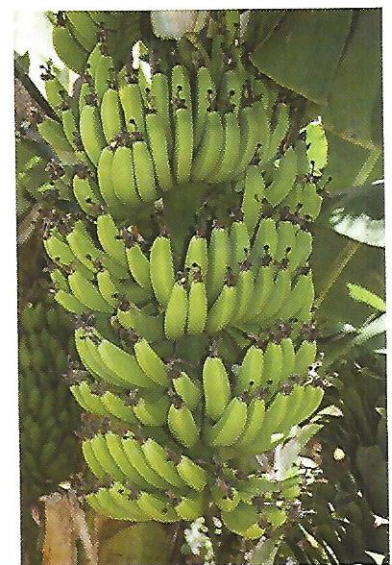
A



B

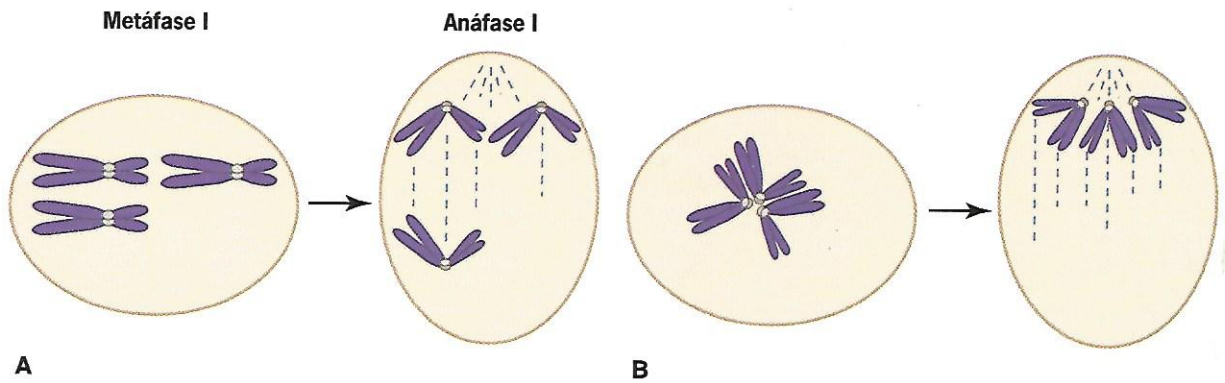


C



D

■ FIGURA 6.7 Vegetais poliploides de significado agrícola ou hortícola: A. crisântemo (tetraploide), B. morango (octaploide), C. algodão (tetraploide), D. banana (triploide).

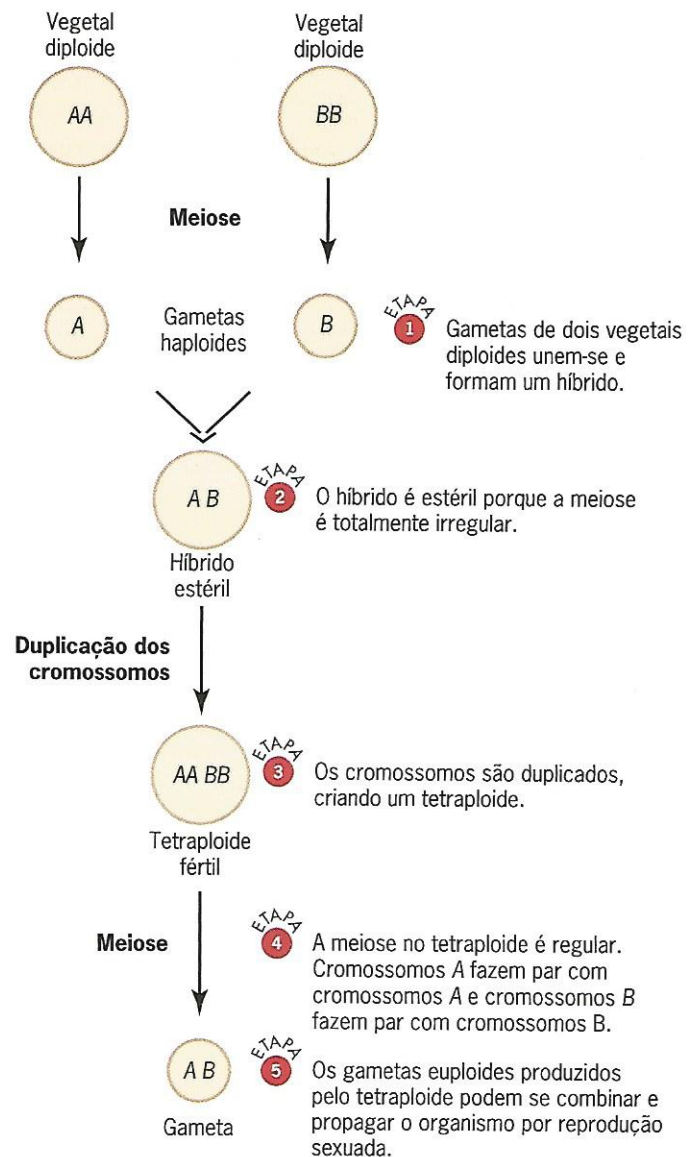


■ FIGURA 6.8 Meiose em organismo triploide. A. Formação univalente. Sinapse de dois dos três homólogos, deixando um univalente livre para se deslocar até um dos polos durante a anáfase. B. Formação trivalente. Sinapse dos três homólogos, formando um trivalente, que pode se deslocar em bloco até um polo durante a anáfase. No entanto, há outras disjunções possíveis na anáfase.

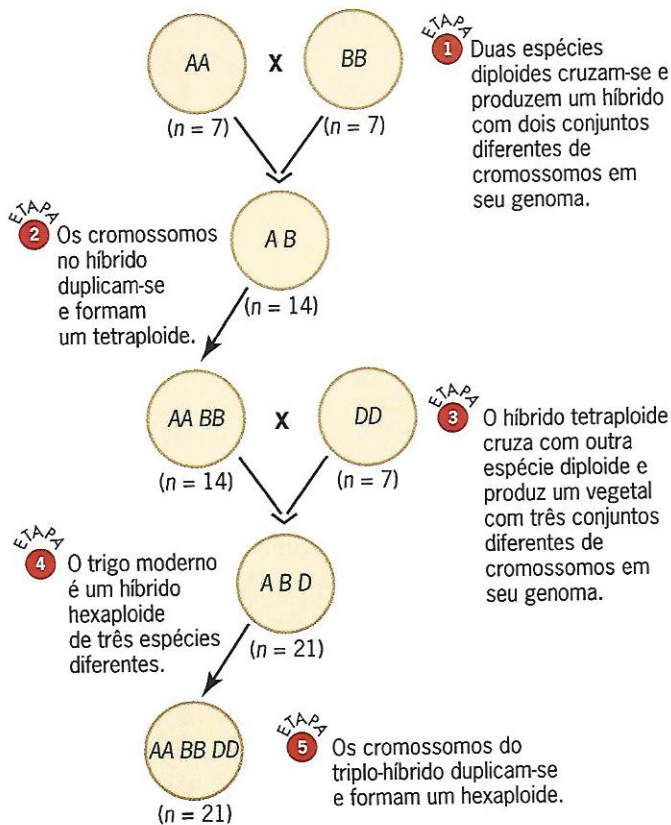
contêm dois conjuntos distintos de cromossomos e que cada conjunto foi duplicado. Assim, os tetraploides férteis parecem ter se originado por duplicação cromossômica em híbrido produzido pelo cruzamento de duas espécies diploides diferentes, mas aparentadas; na maioria das vezes, essas espécies têm números de cromossomos iguais ou muito semelhantes. A Figura 6.9 mostra um mecanismo plausível para a origem desse tetraploide. Dois organismos diploides, A e B, são cruzados e produzem um híbrido que recebe um conjunto de cromossomos de cada espécie parental. Esse híbrido provavelmente será estéril porque não é possível o pareamento dos cromossomos A e B. Contudo, se houver duplicação dos cromossomos desse híbrido, a meiose prosseguirá em ordem razoável. Cada cromossomo A e B poderá formar par com um homólogo perfeito. Assim, a segregação meiótica é capaz de produzir gametas com um conjunto completo de cromossomos A e B. Na fertilização, há união desses gametas “diploides” e formação de zigotos tetraploides, que sobrevivem em razão do equilíbrio de cada conjunto parental de cromossomos.

Evidentemente, essa situação de hibridização entre espécies diferentes, mas relacionadas, seguida por duplicação dos cromossomos ocorreu muitas vezes durante a evolução dos vegetais. Em alguns casos, o processo ocorreu repetidas vezes, gerando poliploides complexos com diferentes conjuntos de cromossomos. Um dos melhores exemplos é o trigo moderno, *Triticum aestivum* (Figura 6.10). Essa importante espécie agrícola é um hexaploide que contém três diferentes conjuntos de cromossomos, todos duplicados. Cada conjunto tem sete cromossomos, com um total de 21 nos gametas e 42 nas células somáticas. Assim, como observamos no início deste capítulo, o trigo moderno parece ter sido formado por dois eventos de hibridização. O primeiro foi a combinação de duas espécies diploides com formação de um organismo tetraploide, e o segundo foi a combinação desse tetraploide com outro diploide e formação de um hexaploide. Os citogeneticistas identificaram cereais primitivos no Oriente Médio que podem ter participado desse processo evolutivo. Em 2010, grande parte do DNA do genoma do trigo foi sequenciada. Esse genoma é mui-

to grande, corresponde a aproximadamente o quádruplo do tamanho do genoma humano. A análise de todas essas sequências de DNA ajudará a compreender a história evolutiva do trigo.



■ FIGURA 6.9 Origem de um tetraploide fértil por hibridização de dois organismos diploides e subsequente duplicação dos cromossomos.



■ **FIGURA 6.10** Origem do trigo hexaploide por hibridização sequencial de diferentes espécies. Cada hibridização é seguida por duplicação dos cromossomos.

Por ser menor a probabilidade de que cromossomos de espécies diferentes interfiram na segregação um do outro durante a meiose, é muito maior a chance de fertilidade de poliploides originados de hibridizações entre espécies diferentes que a de poliploides originados da duplicação de cromossomos de uma única espécie. Os poliploides criados por hibridização entre diferentes espécies são denominados **alopoliploides** (prefixo de origem grega que significa “outro”); nesses poliploides, os genomas formadores são qualitativamente diferentes. Os poliploides criados por duplicação de cromossomos na mesma espécie são denominados **autopoliploides** (prefixo de origem grega que significa “próprio”); nesses poliploides, houve multiplicação de um genoma para criar conjuntos extras de cromossomos.

A duplicação dos cromossomos é um processo essencial na formação de poliploides. Um possível mecanismo é a entrada da célula em mitose sem que haja citocinese. Essa célula terá o dobro do número habitual de cromossomos. Por meio de divisões subsequentes, pode originar um clone de células poliploides, que contribuiriam para a propagação assexuada do organismo ou para a formação de gametas. Em vegetais, é preciso lembrar que a linhagem germinativa não é separada no início do desenvolvimento, como ocorre em animais. Na verdade, os tecidos reprodutivos só se diferenciam depois de muitos ciclos de divisão celular. Se houve duplicação acidental dos cromossomos durante uma dessas divisões celulares, os tecidos reprodutivos desenvolvidos mais tarde podem ser po-

liploides. Outra possibilidade é a alteração da meiose de maneira a produzir gametas não reduzidos (com o dobro do número normal de cromossomos). Se esses gametas participarem da fertilização, os zigotos serão poliploides. Esses zigotos podem dar origem a organismos maduros que, dependendo da natureza da poliploidia, podem ser capazes de produzir gametas. Para compreender melhor essas possibilidades acompanhe a solução do problema no box Resolva!: Pareamento de cromossomos em poliploides.

POLIPLOIDIA E POLITENIA TECIDO-ESPECÍFICA

Em alguns organismos, determinados tecidos tornam-se poliploides durante o desenvolvimento. A poliploidização provavelmente é uma resposta à necessidade de várias cópias de cada cromossomo e dos genes nele existentes. A **endomitose**, processo que produz células poliploides, requer duplicação do cromossomo, seguida por separação das cromátides-irmãs resultantes. No entanto, por não haver divisão celular associada, ocorre acúmulo de cromossomos extranumerários no núcleo. No fígado e no rim humanos, por exemplo, um ciclo de endomitose produz células tetraploides.

Às vezes, a poliploidização ocorre sem a separação das cromátides-irmãs. Nesses casos, os cromossomos duplicados acumulam-se próximo uns dos outros, formando um feixe de filamentos paralelos alinhados. Os cromossomos resultantes são denominados **politênicos**, termo derivado do grego, que significa “muitos filamentos”. Os exemplos mais espetaculares de cromossomos politênicos são encontrados nas glândulas salivares de larvas de *Drosophila*. Cada cromossomo passa por nove ciclos de replicação,

Resolva!

Pareamento de cromossomos em poliploides

Existem seis cromossomos nos gametas da espécie vegetal A e nove cromossomos nos gametas da espécie vegetal B. O cruzamento entre essas duas espécies produziu híbridos estéreis nos quais não se observou pareamento de cromossomos nas células-mães de microsporo das anteras. O genótipo híbrido A × B, porém, pôde ser propagado de maneira vegetativa por enraizamento de estacas das plantas. Uma dessas estacas deu origem a uma planta robusta e, por acaso, fértil; o exame citológico das células-mães de microsporo dessa planta mostrou 15 bivalentes. Em seguida, fez-se o retrocruzamento dessa planta fértil com a espécie A e o exame citológico das células-mães de microsporo da prole. (a) Explique a origem da planta robusta e fértil. (b) Quantos bivalentes você esperaria encontrar nas células-mães de microsporo da prole do retrocruzamento? (c) Quantos cromossomos sem par (univalentes) você esperaria ver nessa prole?

► **Leia a resposta do problema no site**
<http://gen-io.grupogen.com.br>.

produzindo um total aproximado de 500 cópias em cada célula. Há pareamento compacto de todas as cópias com a formação de um feixe espesso de fibras de cromatina. Esse feixe é tão grande que pode ser visto sob pequeno aumento ao microscópio de dissecação. A diferença de espiralamento ao longo do comprimento do feixe causa variação na densidade da cromatina. Quando se aplicam corantes a esses cromossomos, a cor é mais intensa na cromatina mais densa, criando um padrão de faixas claras e escuras (Figura 6.11). Esse padrão tem alta reprodutibilidade, permitindo a análise detalhada da estrutura do cromossomo.

Os cromossomos politênicos de *Drosophila* apresentam duas outras características:

1. Pareamento de cromossomos politênicos homólogos. Normalmente, pensamos no pareamento como uma propriedade dos cromossomos meióticos; mas também há pareamento dos cromossomos somáticos em muitas espécies de insetos, provavelmente um recurso para organizar os cromossomos no núcleo. Quando há pareamento dos cromossomos politênicos de *Drosophila*, os grandes feixes de cromatina tornam-se ainda maiores. Como esse pareamento é preciso – ponto a ponto ao longo da extensão do cromossomo – o alinhamento dos dois homólogos é perfeito. Desse modo, o alinhamento exato dos padrões de bandejamento de cada um deles torna quase impossível distinguir cada membro de um par.
2. Todos os centrômeros de cromossomos politênicos de *Drosophila* condensam-se em um corpo denominado cromocentro. O material presente de cada lado dos centrômeros também é incluído nessa massa. O resultado é que os braços do cromossomo parecem sair do cromocentro. Esses braços, divididos em faixas, consistem em eucromatina, a porção do cromossomo que contém a maioria dos genes; o cromocentro é constituído de heterocromatina, material com poucos genes que circunda o centrômero. Ao contrário dos braços eucromáticos do cromossomo, essa heterocromatina cêntrica não se torna politênica. Assim, replica-se muito menos que a eucromatina.

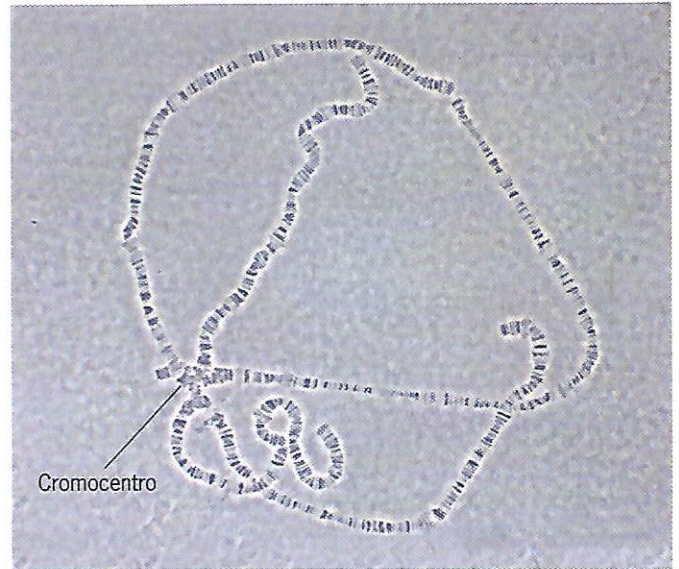


FIGURA 6.11 Cromossomos politênicos de *Drosophila*.

Na década de 1930, C. B. Bridges publicou desenhos detalhados dos cromossomos politênicos (Figura 6.12). Ele dividiu arbitrariamente cada cromossomo em seções, que numerou; depois, dividiu cada seção em subseções, designadas pelas letras A a F. Dentro de cada subseção, Bridges enumerou todas as faixas escuras, criando um repertório alfanumérico de sítios ao longo do comprimento de cada cromossomo. O sistema alfanumérico de Bridges ainda é usado atualmente para descrever as características desses cromossomos excepcionais.

Os cromossomos politênicos de *Drosophila* são retidos na intérfase do ciclo celular. Assim, embora a maioria das análises citológicas seja feita em cromossomos mitóticos, as análises mais completas e detalhadas são realizadas em cromossomos interfásicos politenizados. Esses cromossomos são encontrados em muitas espécies da ordem de insetos Diptera, inclusive em moscas e mosquitos. Infelizmente, os seres humanos não têm cromossomos politênicos; assim, a análise citológica de alta resolução que é possível em *Drosophila* não é possível em nossa própria espécie.

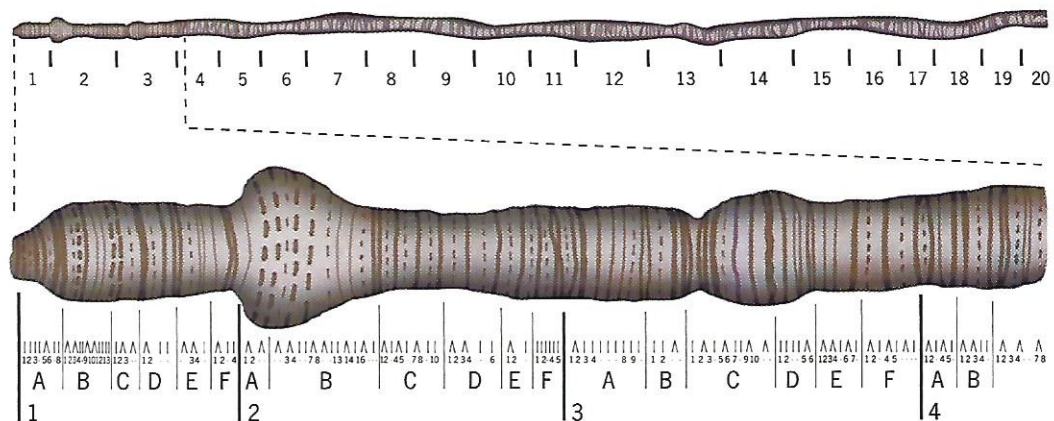


FIGURA 6.12 Mapas de cromossomo politênico de Bridges. (Em cima) Padrão de bandejamento do cromossomo X politênico. O cromossomo é dividido em 20 seções numeradas. (Embaixo) Imagem detalhada da extremidade esquerda do cromossomo X politênico mostrando o sistema de Bridges para designar cada banda.

PONTOS ESSENCIAIS

- Os poliploides contêm conjuntos extras de cromossomos
- Muitos poliploides são estéreis porque a segregação dos vários conjuntos de cromossomos na meiose é irregular
- Os poliploides produzidos por duplicação cromossômica em híbridos interespecíficos podem ser férteis caso haja segregação independente de seus genomas constituintes
- Em alguns tecidos somáticos – por exemplo, as glândulas salivares de larvas de *Drosophila* – há ciclos sucessivos de replicação dos cromossomos sem divisões celulares interpostas, com a produção de grandes cromossomos politênicos que são ideais para análise citogenética.

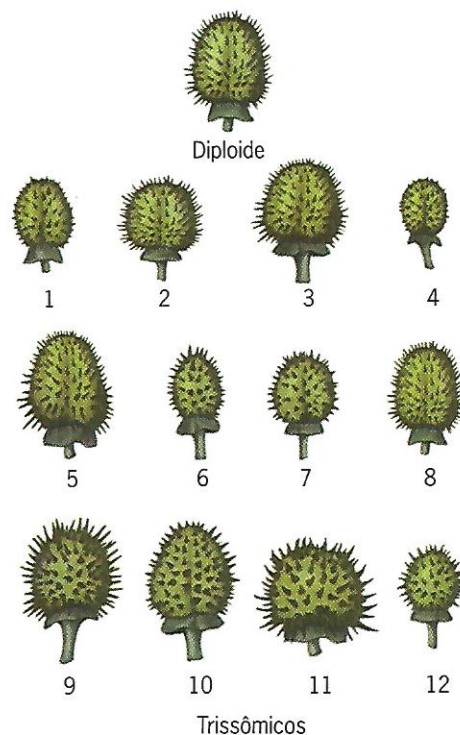
Aneuploidia

A representação insuficiente ou excessiva de um cromossomo ou um segmento de cromossomo pode afetar o fenótipo.

A **aneuploidia** é a alteração numérica de parte do genoma, geralmente a alteração na dose de um único cromossomo. Indivíduos que têm um cromossomo a mais, um cromossomo a menos ou uma combinação dessas anomalias são aneuploides. Essa definição também inclui segmentos de cromossomos. Assim, um indivíduo com deleção do braço de um cromossomo também é considerado aneuploide.

A aneuploidia foi estudada originalmente em vegetais, nos quais se demonstrou que o desequilíbrio cromossômico geralmente tem efeito fenotípico. O estudo clássico foi o de Albert Blakeslee e John Belling, que analisaram anomalias cromossômicas de estramônio, *Datura stramonium*. Essa espécie diploide tem 12 pares de cromossomos, com um total de 24 cromossomos nas células somáticas. Blakeslee colheu plantas com fenótipo alterado e descobriu que em alguns casos a herança dos fenótipos era irregular. Aparentemente, esses mutantes peculiares eram causados por fatores dominantes transmitidos principalmente pela planta do sexo feminino. Examinando os cromossomos das plantas mutantes, Belling constatou a presença de um cromossomo extranumerário em todos os casos. A análise detalhada verificou que o cromossomo extra era diferente em cada linhagem mutante. Ao todo, eram 12 mutantes diferentes, cada um deles correspondente à triplicação de um dos cromossomos de *Datura* (Figura 6.13). Essas triplicações são chamadas **trissomias**. As irregularidades de transmissão desses mutantes eram consequência do comportamento anômalo do cromossomo durante a meiose.

Belling também descobriu o motivo da transmissão preferencial dos fenótipos trissômicos pelo sexo feminino. Durante o crescimento do tubo polínico, o pólen aneuploide – em especial, o pólen com $n + 1$ cromossomos – não compete bem com o pólen euploide. Desse modo, vegetais trissômicos quase sempre herdam o cromossomo extra do genitor de sexo feminino. O trabalho de Belling com *Datura* demonstrou a necessidade de que cada cromossomo esteja presente na dose apropriada para o crescimento e o desenvolvimento normais.

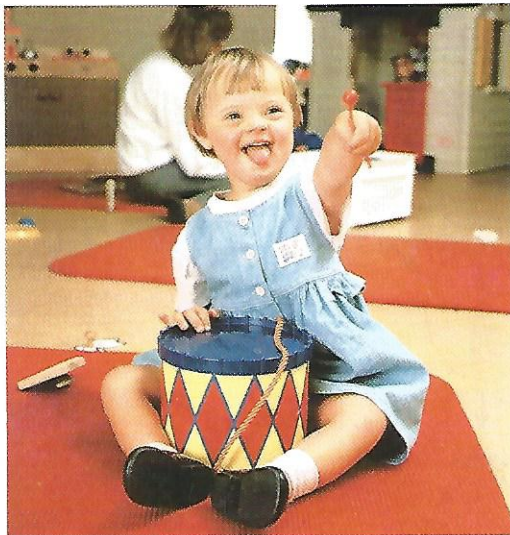


■ FIGURA 6.13 Cápsulas de semente de *Datura stramonium* normal e trissômico. A figura mostra todas as 12 trissomias.

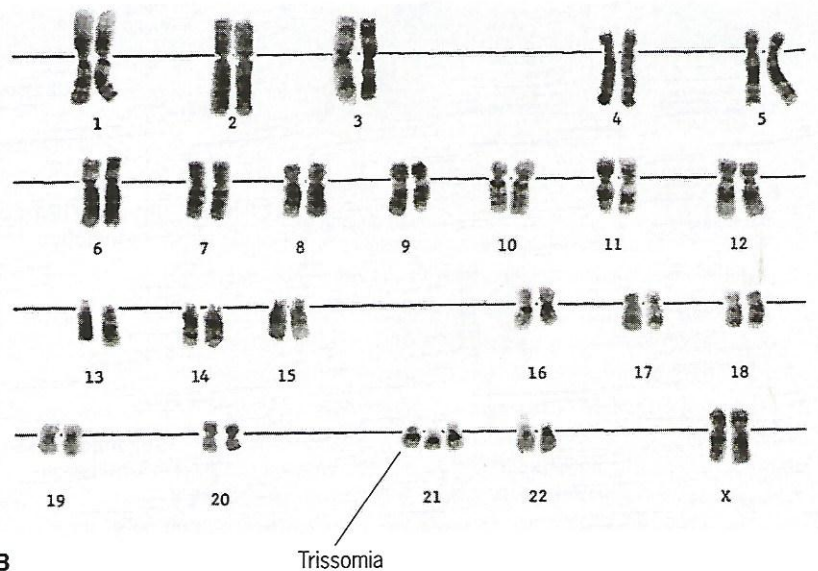
Desde o trabalho de Belling, identificaram-se aneuploides em muitas espécies, inclusive na nossa. Um organismo com ausência de um cromossomo, ou um segmento de cromossomo, é **hipoploide** (prefixo grego que significa “abaixo”). Um organismo com um cromossomo, ou um segmento de cromossomo, a mais é **hiperploide** (prefixo grego que significa “acima”). Esses termos abrangem uma grande variedade de anormalidades.

TRISSOMIA EM SERES HUMANOS

A anormalidade cromossômica mais conhecida e mais comum em seres humanos é a **síndrome de Down**, distúrbio causado por um cromossomo 21 extranumerário (Figura 6.14A). Essa síndrome foi descrita pela primeira vez em 1866, por Langdon Down, médico britânico, mas a base cromossômica só foi compreendida com clareza



A



B

■ FIGURA 6.14 Síndrome de Down. A. Menina com síndrome de Down. B. Cariótipo de uma criança com síndrome de Down, mostrando trissomia do cromossomo 21 (47, XX, +21).

em 1959. As pessoas com síndrome de Down geralmente são baixas e têm hipermobilidade articular, sobretudo nos tornozelos; o crânio é largo, as narinas são amplas, a língua é grande e tem sulcos característicos e as mãos são curtas e largas com uma prega palmar. O comprometimento mental requer orientação e cuidados especiais. O período de vida das pessoas com síndrome de Down é muito menor que das outras pessoas. Além disso, quase sempre elas desenvolvem doença de Alzheimer, um tipo de demência bastante comum em idosos. As pessoas com síndrome de Down, porém, desenvolvem essa doença na quarta ou quinta década de vida, muito mais cedo que as outras pessoas.

O cromossomo 21 é mais na síndrome de Down é um exemplo de trissomia. A Figura 6.14B mostra o cariótipo de uma paciente com síndrome de Down. Existem ao todo 47 cromossomos, entre eles dois cromossomos X além do cromossomo 21 extra. Portanto, o cariótipo é 47, XX, +21.

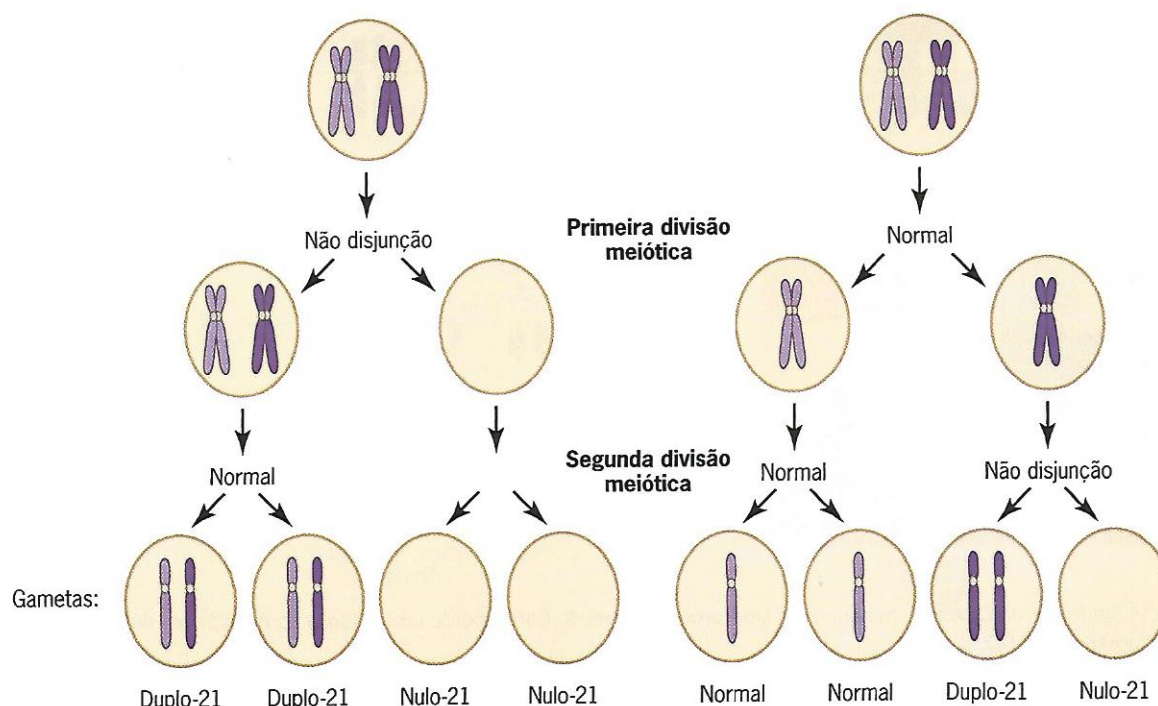
A trissomia do 21 pode ser causada por não disjunção do cromossomo em uma das divisões meióticas (Figura 6.15). O evento de não disjunção pode ocorrer em qualquer um dos pais, porém é mais provável no sexo feminino. Além disso, a frequência de não disjunção aumenta com a idade materna. Assim, nas mulheres com menos de 25 anos, o risco de ter um filho com síndrome de Down é de aproximadamente 1 em 1.500, enquanto nas mulheres de 40 anos, é de 1 em 100. Esse aumento do risco é causado por fatores que afetam adversamente o comportamento meiótico do cromossomo à medida que a mulher envelhece. Nas mulheres, a meiose começa na vida fetal, mas só é concluída depois da fertilização do ovócito. Durante o longo período antes da fertilização, as células meióticas permanecem na prófase da primeira divisão. Nesse estado de pausa, os cromossomos podem perder o par. Quanto maior é a duração da prófase, maior é a chance de que não haja pareamento nem

disjunção subsequente do cromossomo. Portanto, as mulheres mais velhas são mais propensas a produzir ovócitos aneuploides.

Também há relato de trissomias dos cromossomos 13 e 18. No entanto, são raras, e os indivíduos afetados apresentam anormalidades fenotípicas graves e vivem pouco, geralmente morrendo nas primeiras semanas de vida. Outra trissomia viável observada em seres humanos é o cariótipo triplo-X, 47, XXX. Esses indivíduos sobrevivem porque dois dos três cromossomos X são inativados, reduzindo a dose do cromossomo X de maneira que se aproxime do nível normal de um. Os indivíduos triplos X são do sexo feminino e têm fenótipo normal, ou quase; às vezes há leve comprometimento mental e diminuição da fertilidade.

O cariótipo 47, XXY também é uma trissomia viável em seres humanos. Esses indivíduos têm três cromossomos sexuais, dois X e um Y. O fenótipo é masculino, mas também podem apresentar algumas características sexuais secundárias femininas e geralmente são estéreis. Em 1942, H. F. Klinefelter descreveu as anormalidades associadas a esse distúrbio, agora denominado **síndrome de Klinefelter**; inclui testículos pequenos, mamas aumentadas, membros longos, ginecomastia e menor desenvolvimento dos pelos corporais. O cariótipo XXY pode originar-se pela fertilização de um ovócito excepcional XX por um espermatozoide Y ou pela fertilização de um ovócito X por um espermatozoide excepcional XY. O cariótipo XXY representa cerca de três quartos de todos os casos de síndrome de Klinefelter. Outros casos têm cariótipos mais complexos, como XXYY, XXXY, XXXYY, XXXXY, XXXXXY e XXXXXY. Todos os indivíduos com síndrome de Klinefelter têm um ou mais corpúsculos de Barr nas células, e aqueles que têm dois cromossomos X geralmente têm algum grau de comprometimento mental.

O cariótipo 47, XYY é outra trissomia viável em seres humanos. Esses indivíduos são do sexo masculino e, ex-



■ FIGURA 6.15 A não disjunção meiótica do cromossomo 21 e a origem da síndrome de Down. A não disjunção na meiose I produz gametas anormais, que ou têm duas cópias do cromossomo 21 (duplo-21) ou não têm nenhuma cópia desse cromossomo (nulo-21). A não disjunção na meiose II produz um gameta com dois cromossomos-irmãos idênticos (duplo-21) e um gameta sem cromossomo 21 (nulo-21).

ceto pela tendência a serem mais altos que os homens 46, XY, não apresentam uma síndrome constante de características. Todas as outras trissomias em seres humanos são letais no período embrionário, mostrando a importância da dose correta do gene. Ao contrário do que ocorre em *Datura*, no qual todas as trissomias possíveis são viáveis, os seres humanos não toleram muitos tipos de desequilíbrio cromossômico (Tabela 6.1).

MONOSSOMIA

A **monossomia** ocorre quando há ausência de um cromossomo em indivíduo diploide. Em seres humanos, só existe um monossômico viável, o cariótipo 45, X. Esses indivíduos têm um só cromossomo X e um complemento diploide de autossomos. O fenótipo é feminino, mas, por terem ovários rudimentares, são quase sempre estéreis. Os indivíduos 45, X geralmente são baixos; têm pescoço alado, deficiência auditiva e anormalidades cardiovasculares significativas. Henry H. Turner foi o primeiro a descrever o distúrbio em 1938; por isso, agora é denominado **síndrome de Turner**. Os indivíduos 45, X podem originar-se de ovócitos ou espermatozoides sem um cromossomo sexual ou da perda de um cromossomo sexual na mitose algum tempo depois da fertilização (Figura 6.16). Essa última possibilidade é respaldada pela constatação de que muitos indivíduos com síndrome de Turner são *mosaicos somáticos*. Essas pessoas têm dois tipos de células no corpo; algumas são 45, X e outras, 46, XX. Obviamente, esse mosaicismo do cariótipo surge quando há perda de um cromossomo X durante o desenvolvimento de um zigoto 46, XX. Todos os descendentes da célula em que

houve perda são 45, X. Se a perda ocorrer no início do desenvolvimento, uma fração considerável das células do corpo será aneuploide e o indivíduo apresentará características de síndrome de Turner. Se a perda ocorrer mais tarde, a população de células aneuploides será menor, e a intensidade da síndrome tende a ser menor. Veja a análise dos procedimentos usados para detectar aneuploidia em fetos humanos no boxe Em foco: Amniocentese e biopsia de vilosidades coriônicas.

Os mosaicos de cromossomos XX/XO também ocorrem em *Drosophila*, na qual produzem um fenótipo curioso. Como nessa espécie o sexo é determinado pela proporção entre o número de cromossomos X e de autossomos, essas moscas são, em parte, fêmeas e, em parte, machos. As células XX desenvolvem-se na direção feminina, e as células XO desenvolvem-se na direção masculina. Moscas com estruturas masculinas e femininas são **ginandromorfas** (derivado das palavras gregas que significam “mulher”, “homem” e “forma”).

As pessoas com cariótipo 45, X não têm corpúsculos de Barr nas células, indicando que o único cromossomo X presente não foi inativado. Por que, então, as pacientes com síndrome de Turner, que têm o mesmo número de cromossomos X ativos que as mulheres XX normais, têm anormalidades fenotípicas? A resposta provável é que um pequeno número de genes permanece ativo nos dois cromossomos X em mulheres 46, XX normais. Aparentemente, esses genes não inativados são necessários em dose dupla para o crescimento e o desenvolvimento apropriados. A constatação de que pelo menos alguns desses genes especiais ligados ao X também estão presentes no cromossomo Y explicaria por que os homens

Tabela 6.1

Aneuploidia resultante da não disjunção em seres humanos.

Cariótipo	Fórmula cromossômica	Síndrome clínica	Frequência estimada ao nascimento	Fenótipo
47, +21	$2n + 1$	Down	1/700	Mãos largas e curtas com prega palmar, baixa estatura, hiperflexibilidade das articulações, retardo mental, cabeça larga, face redonda, boca aberta com língua grande, prega epicântica.
47, +13	$2n + 1$	Patau	1/20.000	Deficiência mental e surdez, convulsões musculares leves, fenda labial e/ou palatina, anomalias cardíacas, calcanhar proeminente.
47, +18	$2n + 1$	Edward	1/8.000	Malformação congênita de muitos órgãos, orelhas malformadas e de implantação baixa, micrognatia, boca e nariz pequenos com aparência geral de duende, deficiência mental, rim em ferradura ou duplo, esterno curto; 90% morrem nos primeiros 6 meses depois do nascimento.
45, X	$2n - 1$	Turner	1/2.500 recém-nascidos do sexo feminino	Mulher com atraso do desenvolvimento sexual, geralmente estéril, baixa estatura, pescoço alado, anormalidades cardiovasculares, deficiência auditiva.
47, XXY	$2n + 1$	Klinefelter	1/500 recém-nascidos do sexo masculino	Homem subfértil com testículos pequenos, mamas desenvolvidas, voz aguda feminina, genuvalgo, membros longos.
48, XXXY	$2n + 2$			
48, XYY	$2n + 2$			
49, XXXXY	$2n + 3$			
50, XXXXXY	$2n + 4$			
47, XXX	$2n + 1$	Triplo-X	1/700	Mulher com órgãos genitais geralmente normais e fertilidade limitada, retardo mental leve.

XY crescem e se desenvolvem normalmente. Além disso, o cromossomo X que foi inativado nas mulheres 46, XX é reativado durante a oocitogênese, provavelmente porque a função ovariana normal requer duas cópias de alguns genes ligados ao X. Os indivíduos 45, X, que têm só uma cópia desses genes, não satisfazem essa exigência e, portanto, são estéreis.

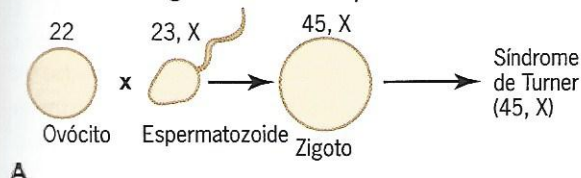
Curiosamente, o análogo do cariótipo de Turner XO no camundongo não causa anormalidades anatômicas.

Esse achado significa que os homólogos no camundongo dos genes humanos implicados na síndrome de Turner precisam estar presentes em apenas uma cópia para o crescimento e o desenvolvimento normais. Para investigar a origem do cariótipo da síndrome de Turner XO, acompanhe o exercício do boxe Problema resolvido: Detecção da não disjunção de cromossomos sexuais.

DELEÇÕES E DUPLICAÇÕES DE SEGMENTOS CROMOSSÔMICOS

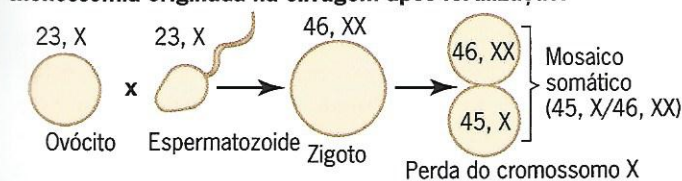
A ausência de um segmento cromossômico é denominada **deleção** ou **deficiência**. Grandes deleções podem ser detectadas citologicamente por estudo dos padrões de bandamento em cromossomos corados, mas as pequenas, não. Em um organismo diploide, a deleção de um segmento cromossômico faz parte do genoma hipoploide. Essa hipoploidia pode estar associada a um efeito fenotípico, principalmente se a deleção for grande. Um exemplo clássico é a síndrome do miado do gato, também conhecida como síndrome *cri-du-chat* (do francês) em seres humanos (Figura 6.17). Esse distúrbio é causado por deleção no braço curto do cromossomo 5. O tamanho da deleção varia. Indivíduos heterozigotos para a deleção e um cromossomo normal têm o cariótipo 46 del(5) (p14), no qual os termos entre parênteses indicam a ausência de bandas na região 14 do braço curto (p) de um dos

Monossomia originada na fertilização.



A

Monossomia originada na clivagem após fertilização.



B

■ FIGURA 6.16 Origem do cariótipo da síndrome de Turner na fertilização (A) ou na clivagem após a fertilização (B).

Em foco

AMNIOCENTESE E BIOPSIA DE VILOSIDADES CORIÔNICAS

O casal Anderson, residente em Minneapolis, esperava o primeiro filho. Nem Donald nem Laura Anderson tinham conhecimento de anormalidades genéticas na família, mas por causa da idade de Laura – 38 anos – decidiram fazer um teste de triagem de aneuploidia fetal. O médico de Laura realizou uma **amniocentese**, procedimento de retirada de uma pequena quantidade de líquido da cavidade ao redor do feto em desenvolvimento por meio de uma agulha introduzida no abdome de Laura (**Figura 1**). Essa cavidade, chamada saco amniótico, é envolvida por uma membrana. Administrou-se anestesia local para evitar o desconforto durante o procedimento. A inserção da agulha foi guiada por ultrassonografia e retirou-se um pouco de líquido amniótico. Como esse líquido contém células nucleadas descamadas do feto, é possível determinar o cariótipo fetal. Em geral, as células fetais são separadas do líquido amniótico por centrifugação e cultivadas durante vários dias a algumas semanas. A análise citológica dessas células mostra se o feto é aneuploide. Outros exames podem ser realizados no líquido colhido do saco amniótico para detectar outros tipos de anormalidades, entre elas os defeitos do tubo neural e alguns tipos de mutação. Os resultados de todos esses exames podem demorar até 3 semanas. No caso de Laura, não foram detectadas anormalidades, e, 20 semanas depois da amniocentese, ela deu à luz uma menina saudável.

A **biopsia de vilosidades coriônicas** é outra técnica para detecção de anormalidades cromossômicas no feto. O cório é uma membrana fetal que se fixa à parede uterina por interdigitações, formando a placenta. As diminutas projeções coriônicas para o tecido uterino são as **vilosidades**. Com 10 a 11 semanas de gestação, antes do desenvolvimento da placenta, pode-se fazer uma biopsia das vilosidades coriônicas com o auxílio de uma cânula plástica introduzida no útero através do colo. A cânula é guiada por ultrassonografia e, quando está em posição, é usada para retirar um pedaço minúsculo de material por aspiração. O material coletado geralmente é uma mistura de tecido materno e fetal. Depois

que esses tecidos são separados por dissecação, é possível analisar as células fetais para pesquisa de anormalidades cromossômicas.

A biopsia de vilosidades coriônicas pode ser feita antes da amniocentese (10 a 11 semanas de gestação *versus* 14 a 16 semanas), mas não é tão confiável. Além disso, parece estar associada a um risco ligeiramente maior de aborto que a amniocentese, talvez 2 a 3%. Por essas razões, tende a ser usada apenas nas gestações em que haja um forte motivo para suspeitar de anormalidade genética. Nas gestações de rotina, como a de Laura Anderson, é preferível fazer a amniocentese.

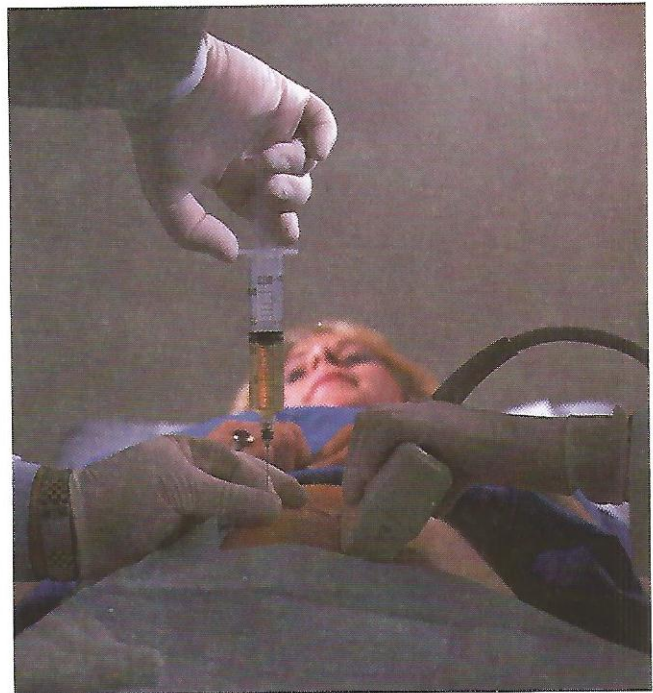


FIGURA 1 O médico colhe uma amostra de líquido do saco amniótico de uma gestante para diagnóstico pré-natal de anormalidade cromossômica ou bioquímica.

cromossomos 5. Esses indivíduos podem apresentar grave comprometimento mental e físico; o choro queixoso, semelhante ao miado de gato na infância, dá nome à síndrome.

A presença de um segmento cromossômico extra é denominada **duplicação**. O segmento extra pode estar unido a um dos cromossomos ou pode constituir um novo cromossomo separado, isto é, uma “duplicação livre”. O efeito é o mesmo nos dois casos: o organismo é **hiperploide** em relação a parte de seu genoma. A exemplo do que ocorre nas deleções, essa hiperploideia pode estar associada a um efeito fenotípico.

As deleções e duplicações são dois tipos de aberrações na estrutura do cromossomo. As grandes aberrações podem ser detectadas por exame dos cromossomos mitóticos corados por agentes de bandeamento como quina- crina ou Giemsa. No entanto, é difícil detectar pequenas

aberrações dessa forma, que geralmente são identificadas por outras técnicas genéticas e moleculares. O organismo mais adequado para estudo de deleções e duplicações é a *Drosophila*, cujos cromossomos politênicos garantem uma oportunidade ímpar de análise citológica detalhada. A **Figura 6.18B** mostra deleção em um dos dois cromossomos homólogos pareados na glândula salivar de *Drosophila*. Em vista da leve separação dos dois cromossomos, é possível perceber a ausência de uma pequena região no cromossomo inferior.

Os segmentos duplicados também podem ser reconhecidos nos cromossomos politênicos. A **Figura 6.18C** mostra duplicação consecutiva (em *tandem*) de um segmento no meio do cromossomo X de *Drosophila*. O pareamento entre as cópias consecutivas desse segmento faz com que pareça haver um nó no meio dos cromossomos. A mutação *Bar* (barra) do olho em *Drosophila* está associa-

PROBLEMA RESOLVIDO



Detecção da não disjunção de cromossomos sexuais

PROBLEMA

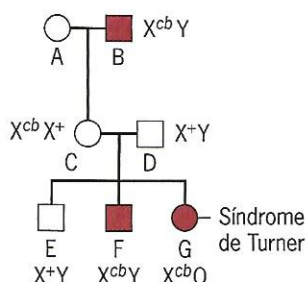
Um homem com discromatopsia casou-se com uma mulher normal. A filha, de fenótipo normal, casou-se com um homem normal e o casal teve três filhos: um menino normal, um menino com discromatopsia e uma menina com discromatopsia e síndrome de Turner. Explique a origem da menina com discromatopsia e síndrome de Turner.

FATOS E CONCEITOS

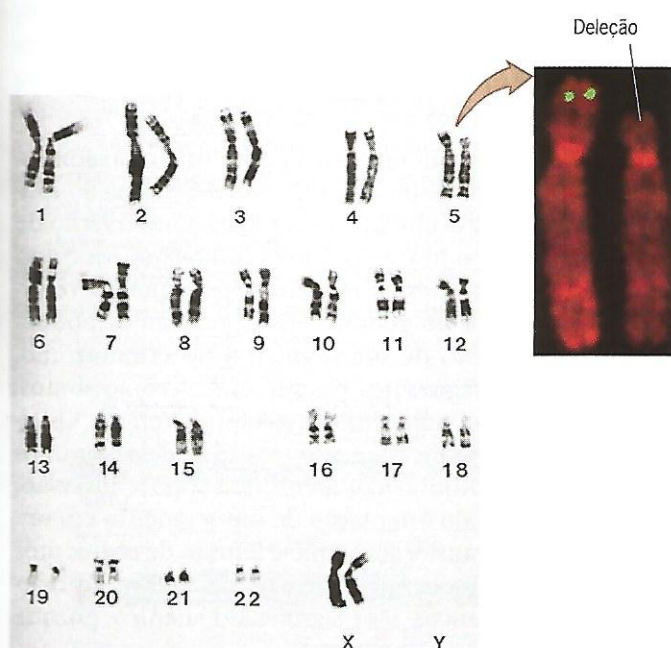
1. A discromatopsia é causada por uma mutação recessiva ligada ao X, cb .
2. A síndrome de Turner é causada por monossomia do cromossomo X (genótipo XO).
3. A monossomia pode ser causada por não disjunção cromossômica durante a mitose ou a meiose.
4. A não disjunção mitótica em indivíduo XX pode criar um mosaico de células XO e XX.

ANÁLISE E SOLUÇÃO

Para iniciar a análise, vamos desenhar o heredograma e identificar todas as pessoas. Além disso, como sabemos que a discromatopsia é causada por mutação recessiva ligada ao X, podemos escrever os genótipos da maioria das pessoas no heredograma.

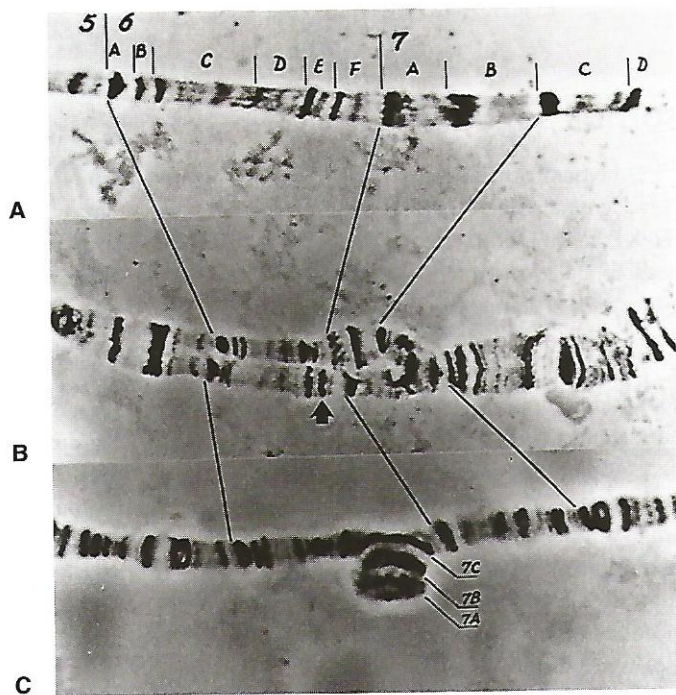


O homem com discromatopsia, B, é uma figura-chave nesse heredograma porque transmitiu obrigatoriamente um cromossomo X com a mutação cb para a filha C, que é a mãe da criança em questão. C não tem discromatopsia, portanto tem de ser heterozigota para o alelo mutante, isto é, seu genótipo é $X^{cb} X^+$. Seu marido, D, também não tem discromatopsia, portanto tem o genótipo $X^+ Y$. Os genótipos dos dois primeiros filhos do casal também são conhecidos com certeza. O último, G, tem síndrome de Turner, o que significa que tem apenas um cromossomo sexual – X. Como essa menina tem discromatopsia, o genótipo provavelmente é X^{cb} O. Esse genótipo poderia ter sido produzido pela fertilização de um ovócito contendo o cromossomo X^{cb} por um espermatozoide sem cromossomo sexual. Nessa situação, não houve disjunção dos cromossomos sexuais durante a meiose no pai de G. Outra possibilidade é que o ovócito que tem X^{cb} tenha sido fertilizado por um espermatozoide com um cromossomo X, que foi perdido durante uma das divisões iniciais do embrião. Nessa segunda hipótese, G seria um mosaico somático de células XO e XX (Figura 6.16B). No entanto, essa explicação não condiz com a observação de que G tem discromatopsia, pois se G fosse um mosaico somático, as células XX teriam de ser $X^{cb} X^+$, e seria esperado que algumas dessas células tivessem formado células fotorreceptoras normais na retina, assim garantindo a visão normal das cores. O fato de G ter discromatopsia indica que ela não tem células $X^{cb} X^+$ na retina – ou provavelmente em nenhuma outra parte do corpo. Portanto, a não disjunção de cromossomos sexuais durante a meiose no pai de G é a explicação mais plausível para o fenótipo de discromatopsia e síndrome de Turner.

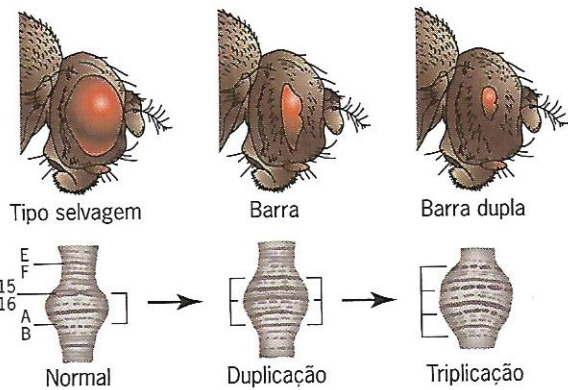


da a uma duplicação consecutiva (Figura 6.19). Essa mutação dominante ligada ao X altera o tamanho e o formato dos olhos compostos, que deixam de ser estruturas esféricas grandes e se transformam em barras estreitas. Na década de 1930, C. B. Bridges analisou cromossomos X com a mutação *Bar* e constatou que a região 16A, que aparentemente continha um gene para formato do olho, havia passado por duplicação consecutiva. Também foram observadas triplicações consecutivas de 16A, e nesses casos o olho composto era pequeníssimo – um fenótipo

■ FIGURA 6.17 Cariótipo de mulher com síndrome *cri-do-chat*, 46 XX del(5)(p14). Há deleção do braço curto de um dos cromossomos 5. O detalhe mostra os dois cromossomos 5 marcados com uma sonda fluorescente gene-específica. O cromossomo à esquerda ligou-se à sonda porque tem esse gene específico, ao passo que o cromossomo à direita não se ligou à sonda porque houve deleção do gene e do material ao seu redor.



■ **FIGURA 6.18** Cromossomos politênicos mostrando (A) a estrutura normal das regiões 6 e 7 no meio do cromossomo X de *Drosophila*, (B) heterozigoto com deleção da região 6F-7C em um dos cromossomos (seta) e (C) um cromossomo X mostrando uma duplicação consecutiva (em *tandem*) invertida da região 6F-7C. Em (B) as bandas proeminentes nas regiões 7A e 7C estão presentes no cromossomo superior, mas ausentes no inferior, indicando que o cromossomo inferior sofreu uma deleção. Em (C) a sequência duplicada é 7C, 7B, 7A, 7A, 7B, 7C da esquerda para a direita.



■ **FIGURA 6.19** Efeitos das duplicações da região 16A do cromossomo X no tamanho dos olhos em *Drosophila*.

denominado barra dupla. Portanto, a intensidade do fenótipo mutante do olho está relacionada com o número de cópias da região 16A – sinal claro da importância da dose do gene na determinação de um fenótipo. Muitas outras duplicações consecutivas foram encontradas em *Drosophila*, na qual a análise de cromossomos politênicos torna a detecção relativamente fácil. Hoje, as técnicas moleculares tornaram possível detectar duplicações consecutivas muito pequenas em uma grande variedade de organismos. Por exemplo, os genes que codificam as proteínas da hemoglobina passaram por duplicação consecutiva em mamíferos (Capítulo 19). As duplicações gênicas parecem ser relativamente comuns e garantem variação significativa para a evolução.

PONTOS ESSENCIAIS

- Nas trissomias, como a síndrome de Down em seres humanos, existem três cópias de um cromossomo; nas monossomias, como a síndrome de Turner em seres humanos, existe apenas uma cópia de um cromossomo
- A aneuploidia pode ser causada por deleção ou duplicação de um segmento cromossômico ou de cromossomos inteiros.

Rearranjos da estrutura do cromossomo

Um cromossomo pode sofrer rearranjo interno ou se unir a outro cromossomo.

Na natureza há considerável variação no número e na estrutura de cromossomos, mesmo entre organismos muito próximos. Por exemplo, a *Drosophila melanogaster* tem quatro pares de cromossomos, que incluem um par de cromossomos sexuais, dois pares de autossomos metacêntricos grandes – cromossomos com o centrômero no meio – e um par de pequenos autossomos puntiformes. A *Drosophila virilis*, cujo parentesco não é muito distante, tem um par de cromossomos sexuais, quatro pares de autossomos acrocêntricos – cromossomos com o centrômero perto de uma extremidade – e um par de autossomos puntiformes. Portanto, as espécies, ainda que do mesmo

gênero, podem ter diferentes arranjos dos cromossomos. Essas diferenças indicam o rearranjo do genoma ao longo do processo de evolução. Na verdade, a observação de que é possível encontrar rearranjos cromossômicos como variantes em uma mesma espécie sugere que há reformulação contínua do genoma. Esses rearranjos podem modificar a posição de um segmento do cromossomo, ou podem unir segmentos de diferentes cromossomos. Em qualquer caso, a ordem dos genes é alterada. Os citogeneticistas identificaram muitos tipos de rearranjos cromossômicos. Aqui analisamos dois tipos: inversão, que é a mudança de orientação de um segmento em um cromossomo, e translocação, que é a fusão de segmentos de diferentes cromossomos. Em seres humanos, os rearranjos cromossômicos têm significado médico porque

alguns predispõem ao desenvolvimento de certos tipos de câncer. Abordamos esses tipos de rearranjos, e sua relação com o câncer, no Capítulo 21.

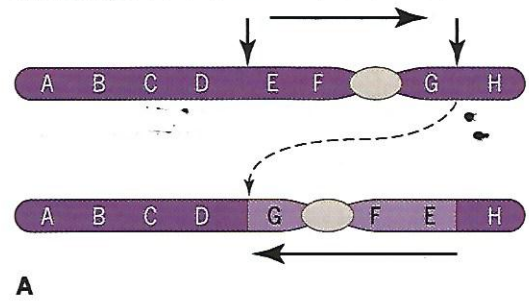
INVERSÕES

A **inversão** ocorre quando um segmento do cromossomo se desprende, gira cerca de 180° e se fixa novamente ao restante do cromossomo; logo, há uma inversão da ordem dos genes no segmento. Esses rearranjos podem ser induzidos no laboratório com raios X, que fragmenta os cromossomos. Às vezes, os fragmentos se fixam novamente, mas durante o processo um segmento gira e há uma inversão. Também há evidências de inversões naturais por atividade de elementos transponíveis – sequências de DNA capazes de mudar de posição no cromossomo (Capítulo 17). Às vezes, durante o deslocamento, esses elementos quebram o cromossomo em fragmentos, que voltam a se unir de maneira anômala, produzindo uma inversão. As inversões também podem ser provocadas pela reunião de fragmentos do cromossomo gerados por cisalhamento mecânico, talvez em razão do entrelaçamento de cromossomos no núcleo. Ninguém sabe ao certo que fração das inversões naturais é causada por cada um desses mecanismos.

Os citogeneticistas distinguem dois tipos de inversões observando se o segmento invertido inclui ou não o centrômero do cromossomo (Figura 6.20). As inversões **pericêntricas** incluem o centrômero e as inversões **paracêntricas**, não. A consequência disso é que uma inversão pericêntrica pode alterar os comprimentos relativos dos dois braços do cromossomo, enquanto a inversão paracêntrica não tem esse efeito. Assim, se um cromossomo acrocêntrico sofre uma inversão com um ponto de quebra em cada braço (*i. e.*, uma inversão pericêntrica), pode ser transformado em um cromossomo metacêntrico. Mas se um cromossomo acrocêntrico sofre uma inversão em que as duas quebras estão no braço longo (*i. e.*, uma inversão paracêntrica), a morfologia do cromossomo não se altera. Desse modo, com o uso de métodos citológicos padronizados, é muito mais fácil detectar inversões pericêntricas que inversões paracêntricas.

Um indivíduo no qual há inversão de um cromossomo, mas não de seu homólogo, é um heterozigoto para inversão. Durante a meiose, há pareamento ponto a ponto dos cromossomos invertido e não invertido ao longo de seu comprimento. Entretanto, por causa da inversão, os cromossomos precisam formar uma alça para permitir o pareamento na região em que a ordem dos genes

Inversão pericêntrica – inclui o centrômero.



Inversão paracêntrica – exclui o centrômero.

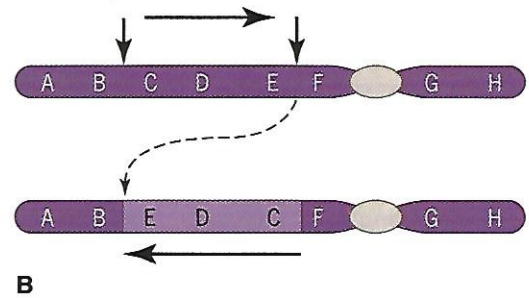


FIGURA 6.20 Inversões pericêntrica e paracêntrica. O cromossomo foi quebrado em dois pontos, e o segmento entre eles foi invertido. Uma inversão pericêntrica (A) modifica o tamanho dos braços do cromossomo porque o centrômero está incluído na inversão. Já na inversão paracêntrica (B) isso não ocorre, porque exclui o centrômero.

foi invertida. A Figura 6.21 mostra essa configuração de pareamento; apenas um cromossomo forma uma alça, e o outro se acomoda ao seu redor. Na prática, a alça para maximizar o pareamento pode surgir no cromossomo invertido ou no cromossomo não invertido. No entanto, perto das extremidades da inversão, os cromossomos são estendidos e há uma tendência à perda da sinapse. Abordamos as consequências genéticas da heterozigidade para inversão no Capítulo 7.

TRANSLOCAÇÕES

A **translocação** ocorre quando um segmento de um cromossomo se desprende e se une a outro cromossomo (*i. e.*, não homólogo). O significado genético é a transferência dos genes de um cromossomo para outro.

A troca de fragmentos de dois cromossomos não homólogos sem perda de material genético é denominada **translocação recíproca**. A Figura 6.22A mostra uma translocação recíproca entre dois grandes autossomos. Três cromossomos trocaram trechos de seus braços direitos.

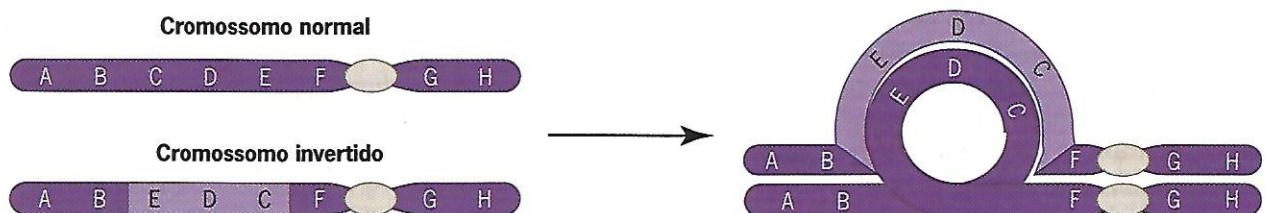
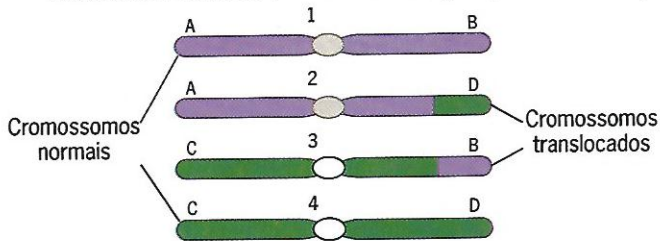


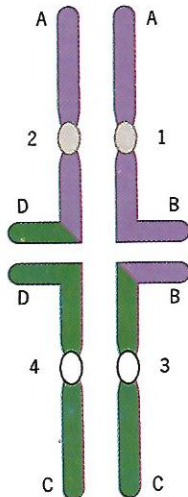
FIGURA 6.21 Pareamento de cromossomos normais e invertidos.

Estrutura de cromossomos no heterozigoto para translocação.



A

Pareamento de cromossomos no heterozigoto para translocação.



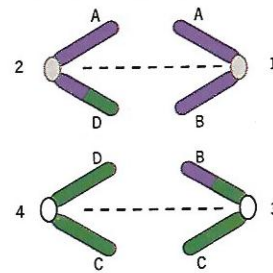
B

■ FIGURA 6.22 Estrutura e comportamento de pareamento de uma translocação recíproca entre cromossomos. Em (B) há pareamento durante a prófase da meiose I, depois da duplicação dos cromossomos.

Durante a meiose, espera-se que haja pareamento cruciforme entre esses cromossomos com translocação e seus homólogos sem translocação (Figura 6.22B). Os dois cromossomos translocados ficam de frente um para o outro no centro da cruz, e os dois cromossomos não translocados fazem o mesmo; para maximizar o pareamento, os cromossomos translocados e não translocados alternam-se, formando os braços da cruz. Essa configuração de pareamento é diagnóstica de um heterozigoto para translocação. As células em que os cromossomos translocados são homozigotos não formam um padrão cruciforme. Em vez disso, há pareamento uniforme de cada cromossomo translocado com seu par de estrutura idêntica.

Como participam do pareamento cruciforme quatro centrômeros, cuja distribuição coordenada para polos opostos na primeira divisão meiótica pode ou não ocorrer, a disjunção do cromossomo em heterozigotos para translocação é um processo um tanto incerto, propenso a produzir gametas aneuploides. Ao todo, há três processos de disjunção possíveis, mostrados na Figura 6.23. Essa figura simplificada só mostra uma das duas cromátides-irmãs de cada cromossomo. Além disso, cada centrômero é identificado para que se possa acompanhar os movimentos do cromossomo; os dois centrômeros brancos são homólogos (*i. e.*, derivados do mesmo par

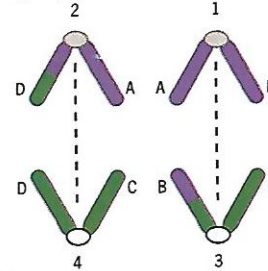
Disjunção adjacente I.



Os centrômeros 1 e 3 vão para um polo e os centrômeros 2 e 4 vão para o outro, produzindo gametas aneuploides.

A

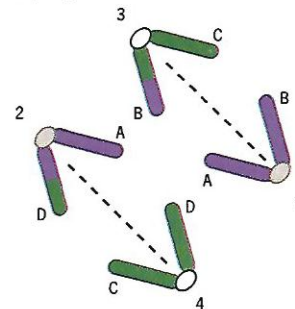
Disjunção adjacente II.



Os centrômeros 1 e 2 vão para um polo e os centrômeros 3 e 4 vão para o outro, produzindo gametas aneuploides.

B

Disjunção alternada.



Os centrômeros 2 e 3 vão para um polo e os centrômeros 1 e 4 vão para o outro, produzindo gametas euploides.

C

■ FIGURA 6.23 Tipos de disjunção em um heterozigoto para translocação durante a meiose I. Para simplificar, é mostrada apenas uma cromátide-irmã de cada cromossomo duplicado. A. Um tipo de disjunção adjacente no qual centrômeros homólogos seguem até polos opostos durante a anáfase. B. Outro tipo de disjunção adjacente no qual centrômeros homólogos seguem para o mesmo polo durante a anáfase. C. Disjunção alternada na qual centrômeros homólogos seguem até polos opostos durante a anáfase.

de cromossomos), do mesmo modo que os dois centrômeros cinza.

Se os centrômeros 2 e 4 seguem para o mesmo polo, forçando o deslocamento de 1 e 3 para o polo oposto, todos os gametas resultantes serão aneuploides, porque alguns segmentos cromossômicos terão deficiência de genes, e outros serão duplicados (Figura 6.23A). Da mesma maneira, se os centrômeros 1 e 2 vão para um polo e os centrômeros 3 e 4 vão para o outro, são produzidos apenas gametas aneuploides (Figura 6.23B). Todos esses casos são denominados *disjunção adjacente* porque os centrômeros que estavam próximos uns dos outros no padrão cruciforme foram para o mesmo polo. Quando os centrômeros que vão para o mesmo polo são de cromossomos diferentes (*i. e.*, *heterólogos*), a disjunção é denominada adjacente I (Figura 6.23A); quando os centrômeros que

vão para o mesmo polo são do mesmo cromossomo (*i. e.*, *homólogos*), a disjunção é denominada adjacente II (Figura 6.23B). Outra possibilidade é a de que os centrômeros 1 e 4 sigam para o mesmo polo, forçando o deslocamento de 2 e 3 para o polo oposto. Esse caso, denominado *disjunção alternada*, produz apenas gametas euploides, embora metade deles tenha apenas cromossomos translocados (Figura 6.23C).

A produção de gametas aneuploides por disjunção adjacente explica por que os heterozigotos para translocação têm fertilidade reduzida. Quando esses gametas fertilizam um gameta euploide, o zigoto resultante é geneticamente desequilibrado e, portanto, é improvável que sobreviva. Em vegetais, muitas vezes os próprios gametas aneuploides são inviáveis, sobretudo do lado masculino, e são produzidos menos zigotos. Portanto, os heterozigotos para translocação são caracterizados por baixa fertilidade. Investigue esse efeito acompanhando o boxe Resolva!: Aborto de pólen em heterozigotos para translocação.

CROMOSSOMOS COMPOSTOS E TRANSLOCAÇÕES ROBERTSONIANAS

Às vezes, um cromossomo funde-se ao seu homólogo, ou duas cromátides-irmãs se unem, formando uma unidade genética única. Um **cromossomo composto** pode ser estável em uma célula desde que tenha apenas um centrômero ativo; se houver dois centrômeros, cada um deles pode se mover para um polo diferente durante a divisão, separando o cromossomo composto. O cromossomo composto também pode ser formado pela união de segmentos de cromossomos homólogos. Por exemplo, os braços direitos dos dois segundos cromossomos em *Drosophila* poderiam se soltar dos braços esquerdos e se fundir no centrômero, criando um meio-cromossomo composto. Às vezes a citogenética chama essa estrutura de **isocromossomo** (do prefixo grego para “igual”), porque seus dois braços são equivalentes. A diferença entre o cromossomo composto e a translocação é que o primeiro é a fusão de segmentos de cromossomos homólogos. Já a translocação é a fusão de cromossomos não homólogos.

O primeiro cromossomo composto foi descoberto em 1922 por Lillian Morgan, casada com T. H. Morgan. Esse cromossomo foi formado pela fusão de dois cromossomos X em *Drosophila*, criando um **cromossomo X ligado** ou X duplo. A descoberta foi feita por meio de experimentos genéticos, não por análise citológica. Lillian Morgan cruzou fêmeas homozigotas para uma mutação recessiva ligada ao X com machos de tipo selvagem. A partir desse cruzamento, seria de esperar que todas as fêmeas da prole fossem do tipo selvagem e os machos, mutantes. No entanto, Morgan observou exatamente o oposto: todas as fêmeas eram mutantes e todos os machos, de tipo selvagem. Outra pesquisa mostrou que os cromossomos X nas fêmeas mutantes haviam se ligado uns aos outros. A Figura 6.24 ilustra o significado genético dessa ligação. As fêmeas com X ligado produzem dois tipos de ovócitos, duplo-X e nulo-X, e os machos produzem dois tipos de espermatozoides, que têm

Resolva!

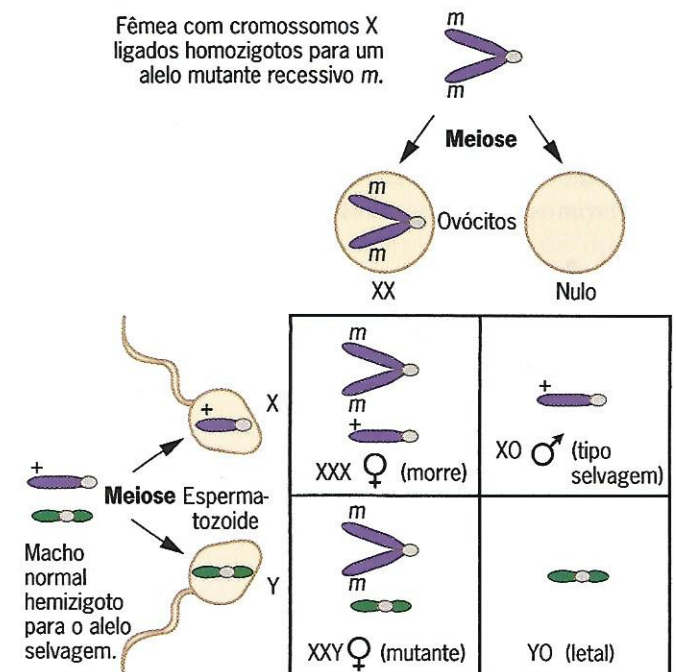
Aborto de pólen em heterozigotos para translocação

Em muitas espécies de vegetais, o pólen aneuploide é inviável. Suponha que um desses vegetais seja heterozigoto para uma translocação recíproca entre dois grandes cromossomos. Se as disjunções adjacente I, adjacente II e alternada nesse heterozigoto para translocação ocorrerem com frequências iguais, que fração do pólen você esperaria que fosse abortada?

► **Leia a resposta do problema no site**
<http://gen-io.grupogen.com.br>.

X e que têm Y. A união desses gametas de todas as maneiras possíveis produz dois tipos de prole viável: fêmeas XXY mutantes, que herdam os cromossomos X ligados da mãe e um cromossomo Y do pai; e os machos XO de fenótipo selvagem, que herdam um cromossomo X do pai e não herdam cromossomo sexual da mãe. Já que o cromossomo Y é necessário para a fertilidade, esses machos XO são estéreis. Lillian Morgan conseguiu propagar os cromossomos X ligados por retrocruzamento de fêmeas XXY com machos XY de tipo selvagem de outro estoque. Por terem herdado um cromossomo Y da mãe, os machos nascidos desse cruzamento eram férteis e puderam ser cruzados com suas irmãs XXY para criar um estoque no qual os cromossomos X ligados foram permanentemente mantidos na linhagem feminina.

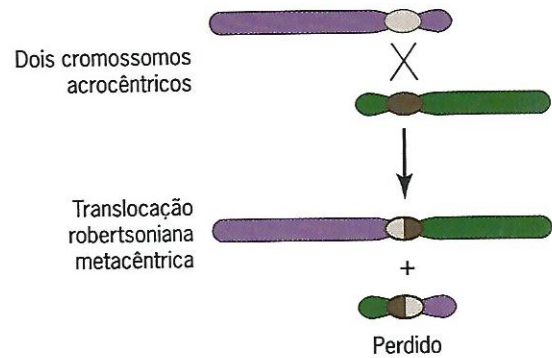
Os cromossomos não homólogos também podem se fundir em seus centrômeros, com criação de uma estrutura chamada **translocação robertsoniana** (Figura 6.25), em homenagem ao citologista F. W. Robertson. Por exemplo,



■ FIGURA 6.24 Resultados de um cruzamento entre um macho normal e uma fêmea com cromossomos X ligados.

a fusão de dois cromossomos acrocêntricos produz um cromossomo metacêntrico; os diminutos braços curtos dos cromossomos participantes são perdidos nesse processo. Ao que tudo indica, essas fusões de cromossomos foram bastante frequentes durante a evolução.

Também pode haver fusão terminoterminal dos cromossomos para formar uma estrutura com dois centrômeros. Se um dos centrômeros for inativado, a fusão dos cromossomos será estável. Evidentemente, essa fusão ocorreu na evolução de nossa própria espécie. O cromossomo 2 humano, que é metacêntrico, tem braços que correspondem a dois cromossomos acrocêntricos nos genomas dos grandes primatas. A análise citológica detalhada mostrou que aparentemente as extremidades dos braços curtos desses dois cromossomos se fundiram para criar o cromossomo 2 humano.



■ FIGURA 6.25 Formação de uma translocação robertsoniana metacêntrica por permuta entre dois cromossomos acrocêntricos não homólogos.

PONTOS ESSENCIAIS

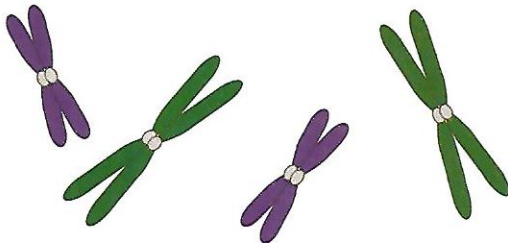
- A inversão faz com que a ordem de genes em um segmento de um cromossomo seja o oposto da ordem inicial
- A translocação permuta segmentos entre dois cromossomos não homólogos
- Os cromossomos compostos são produzidos pela fusão de cromossomos homólogos ou pela fusão dos braços dos cromossomos homólogos
- As translocações robertsonianas resultam da fusão de cromossomos não homólogos.

Exercícios

Aplique a análise genética básica

1. Uma espécie tem dois pares de cromossomos, um longo e outro curto. Desenhe os cromossomos na metáfase da mitose. Indique cada cromátide. Os cromossomos homólogos estão pareados?

Resposta: A metáfase mitótica nessa espécie seria semelhante à mostrada na figura. Como cada cromossomo é duplicado, ele é constituído de duas cromátides-irmãs. No entanto, como a figura mostra a mitose, e não a meiose, não há pareamento dos cromossomos homólogos.



2. A espécie vegetal A apresenta 10 bivalentes de cromossomos na metáfase da meiose I; a espécie vegetal B apresenta 14 bivalentes nesse estágio. As duas espécies são cruzadas, e os cromossomos na prole são duplicados. (a) Quantos bivalentes serão

observados na metáfase da meiose I na prole? (b) Espera-se que a prole seja fértil ou estéril?

Resposta: (a) A prole é uma combinação dos cromossomos dos dois pais. Na espécie A, o número básico de cromossomos é 10; na espécie B, é 14. Portanto, o número básico de cromossomos na prole é $10 + 14 = 24$, e, com os cromossomos duplicados, esse é o número de bivalentes que deve ser observado na metáfase da meiose I. (b) A prole é um alotetraploide e, portanto, deve ser fértil.

3. Quais são os cariótipos de (a) uma mulher com síndrome de Down, (b) um homem com trissomia do 13, (c) uma mulher com síndrome de Turner, (d) um homem com síndrome de Klinefelter, (e) um homem com deleção no braço curto do cromossomo 11?

Resposta: (a) 47, XX, +21, (b) 47, XY, +13, (c) 45, X, (d) 47, XXY, (e) 46, XY del(11p).

4. Que tipo de configuração de pareamento seria observada na prófase da meiose I em (a) um heterozigoto para inversão, (b) um heterozigoto para translocação?

Resposta: (a) Configuração em alça, (b) configuração em cruz.

Autoavaliação

Integre diferentes conceitos e técnicas

- Uma geneticista de *Drosophila* obteve fêmeas com cromossomos X ligados homozigotos para uma mutação recessiva (*y*) que torna o corpo amarelo em vez de cinza. Em um experimento, ela cruza algumas dessas fêmeas com machos de tipo selvagem comuns, e em outro, ela cruza essas fêmeas com machos de tipo selvagem que têm cromossomos X e Y ligados; isto é, têm um cromossomo XY composto. Preveja os fenótipos da prole desses dois cruzamentos e indique o fenótipo estéril, se houver algum.

Resposta: Para prever os fenótipos da prole, precisamos conhecer os genótipos. A maneira mais fácil de determinar esses genótipos é criar um diagrama com os tipos de zigotos produzidos por cada cruzamento.

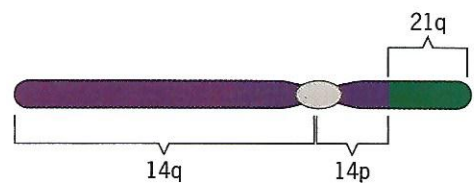
Primeiro, analisamos o cruzamento entre fêmeas com X ligado e corpo amarelo e os machos de tipo selvagem comuns. As fêmeas produzem dois tipos de gametas, X e nulo. Os machos também produzem dois tipos de gametas, X e Y. A combinação de todas as maneiras possíveis produz quatro tipos de zigotos; no entanto, apenas dois tipos são viáveis. Os zigotos XXY dão origem a fêmeas de corpos amarelos – como suas mães, a não ser pelo fato de terem um cromossomo X – e os zigotos XO dão origem a machos de corpo cinza – como seus pais, exceto pela ausência de um cromossomo Y. O cromossomo Y extra nas fêmeas não tem efeito sobre a fertilidade, mas o cromossomo Y ausente nos machos causa esterilidade.

		Ovócitos	
		$\widehat{X^yX^y}$	0
Espermatozoides	X^+	$\widehat{X^yX^y} X^+$ (morrem)	X^+0 machos cinza
	Y	$\widehat{X^yX^y} Y$ fêmeas amarelas	YO (morrem)

Agora consideramos o cruzamento entre fêmeas com X ligado e corpo amarelo e os machos com cromossomo XY composto. Ambos os sexos produzem dois tipos de gametas – as fêmeas produzem os mesmos tipos acima, e os machos produzem XY ou nulo. Quando esses são unidos de todas as maneiras possíveis, observamos que dois tipos de zigotos são viáveis: fêmeas de corpo amarelo com cromossomos X ligados e machos de corpo cinza com um cromossomo XY composto. Os dois tipos dessa prole viável são férteis.

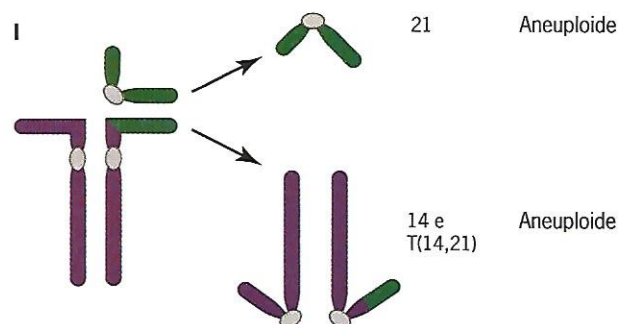
		Ovócitos	
		$\widehat{X^y} \widehat{X^y}$	0
Espermatozoides	$\widehat{X^+Y}$	$\widehat{X^y} \widehat{X^y} \widehat{X^+Y}$ (morrem)	$\widehat{X^+Y} 0$ machos cinza
	0	$\widehat{X^y} \widehat{X^y} 0$ fêmeas amarelas	$0 0$ (morrem)

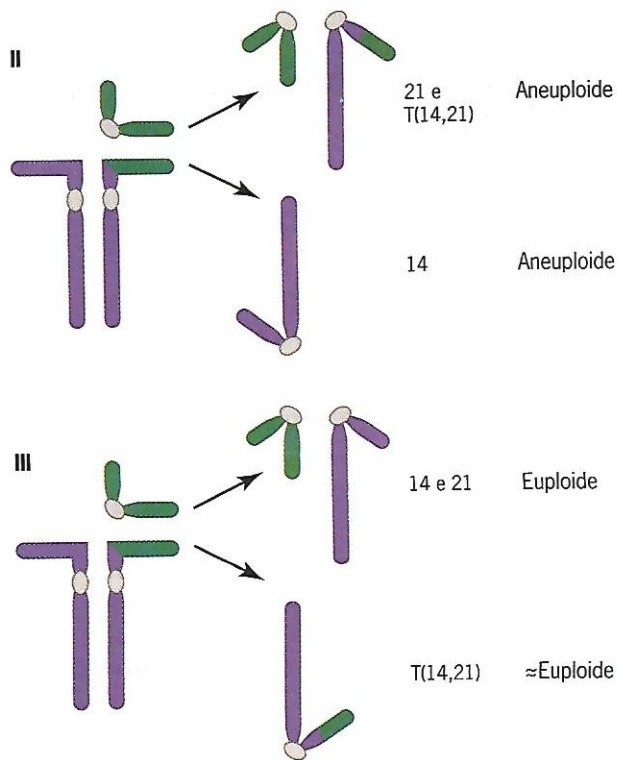
- Um homem com fenótipo normal tem um cromossomo translocado que contém todo o braço longo do cromossomo 14, parte do braço curto do cromossomo 14 e a maior parte do braço longo do cromossomo 21:



O homem também tem um cromossomo 14 normal e um cromossomo 21 normal. Se ele se casar com uma mulher normal do ponto de vista citológico (e fenotípico), há alguma chance de que o casal tenha um filho com fenótipo anormal?

Resposta: Sim, o casal teria filhos com síndrome de Down em consequência da segregação meiótica no homem citologicamente anormal. Durante a meiose nesse homem, o cromossomo translocado, T(14, 21), fará sinapse com os cromossomos normais 14 e 21, formando um trivalente. A disjunção desse trivalente produz seis tipos diferentes de espermatozoides, quatro deles aneuploides.





A fertilização de um ovócito contendo um cromossomo 14 e um cromossomo 21 por qualquer espermatozoide aneuploide produz um zigoto aneuploide, como mostra a tabela associada. Embora a trissomia ou monossomia do cromossomo 14 e a monossomia do cromossomo 21 sejam condições letais, a trissomia do cromossomo 21 não é. Assim, é possível que o casal tenha um filho com síndrome de Down.

Disjunção	Espermatózoide	Zigoto	Distúrbio	Resultado
I	21	14, 21, 21	monossomia do 14	morre
	14, T(14, 21)	14, 14, T(14, 21), 21	trissomia do 14	morre
II	14	14, 14, 21	monossomia do 21	morre
	T(14, 21), 21	14, T(14, 21), 21, 21	trissomia do 21	Down
III	14, 21	14, 14, 21, 21	euploide	normal
	T(14, 21)	14, T(14, 21), 21	≈euploide	normal

Avaliação adicional

Entenda melhor e desenvolva a capacidade analítica

- 6.1 No cariótipo humano, o cromossomo X tem aproximadamente o mesmo tamanho de sete dos autosomos (denominado grupo C de cromossomos). Que procedimento poderia ser usado para distinguir o cromossomo X dos outros membros desse grupo?
- 6.2 Em seres humanos, um cromossomo 22 citologicamente anormal, denominado cromossomo Philadelphia, nome da cidade em que foi descoberto, está associado à leucemia crônica. Esse cromossomo perdeu parte do braço longo. Como você indicaria o cariótipo de um indivíduo que tinha 46 cromossomos nas células somáticas, entre eles um cromossomo 22 normal e um cromossomo Philadelphia?
- 6.3 Durante a meiose, por que alguns tetraploides se comportam de maneira mais regular que os triploides?
- 6.4 A tabela a seguir apresenta dados dos cromossomos de quatro espécies de vegetais e dos híbridos da F_1 :

Espécie ou híbrido da F_1	Metáfase da meiose I		
	Número de cromossomos na extremidade da raiz	Número de bivalentes	Número de univalentes
A	20	10	0
B	20	10	0

- | | | | |
|-------|----|---|----|
| C | 10 | 5 | 0 |
| D | 10 | 5 | 0 |
| A × B | 20 | 0 | 20 |
| A × C | 15 | 5 | 5 |
| A × D | 15 | 5 | 5 |
| C × D | 10 | 0 | 10 |
- (a) Deduzir a origem cromossômica da espécie A.
- (b) Quantos bivalentes e univalentes você esperaria observar na metáfase da meiose I em um híbrido das espécies C e B?
- (c) Quantos bivalentes e univalentes você esperaria observar na metáfase da meiose I em um híbrido das espécies D e B?
- 6.5 Uma espécie de vegetal A, cujos gametas têm sete cromossomos, foi cruzada com uma espécie B aparentada, que tem nove. Os híbridos foram estéreis, e a observação microscópica das células-mães de pólen não mostrou pareamento de cromossomos. Uma parte de um dos híbridos de crescimento vigoroso foi propagada vegetativamente, com a produção de um vegetal que tinha 32 cromossomos em suas células somáticas. Essa planta era fértil. Explique.
- 6.6 Uma espécie de vegetal X com $n = 5$ foi cruzada com uma espécie aparentada Y com $n = 7$. O híbrido da F_1 produziu apenas alguns grãos de pólen, que foram usados para fertilizar as oosferas da espécie Y. Esse cruzamento produziu algumas

plantas, todas com 19 cromossomos. Depois da autofertilização, os híbridos da F_1 produziram algumas plantas da F_2 , cada uma delas com 24 cromossomos. Essas plantas tinham fenótipo diferente das duas espécies originais e eram muito férteis. Explique a sequência de ocorrências que produziu esses híbridos férteis da F_2 .

- 6.7 Identifique os fenótipos sexuais dos seguintes genótipos em seres humanos: XX, XY, XO, XXX, XXY, XYY
- 6.8 Se houver não disjunção do cromossomo 21 na divisão de um ovócito secundário em uma mulher, qual é a chance de que um ovócito maduro derivado dessa divisão receba dois cromossomos 21?
- 6.9 Uma fêmea de *Drosophila* homozigota para uma mutação recessiva ligada ao X causadora de corpo amarelo foi cruzada com um macho de tipo selvagem. Na prole, uma mosca tinha áreas de pigmento amarelo em um corpo cinza. Essas áreas amarelas eram distintamente masculinas, enquanto as áreas cinza eram femininas. Explique o fenótipo peculiar dessa mosca.
- 6.10 O quarto cromossomo de *Drosophila* é tão pequeno que as moscas monossômicas ou trissômicas para ele sobrevivem e são férteis. Vários genes, entre eles o *eyeless* (*ey*), foram localizados nesse cromossomo. Se uma mosca citologicamente normal homozigota para uma mutação *eyeless* (sem olhos) recessiva for cruzada com uma mosca monossômica para um quarto cromossomo de tipo selvagem, quais serão os tipos de prole produzidos? Em que proporções?
- 6.11 Uma mulher com discromatopsia ligada ao X e síndrome de Turner tinha um pai com discromatopsia e mãe normal. A não disjunção dos cromossomos sexuais ocorreu no pai ou na mãe dessa mulher?
- 6.12 Em seres humanos, a síndrome de Hunter é reconhecidamente uma característica ligada ao X com penetrância completa. Na família A, dois indivíduos de fenótipo normal tiveram um filho normal, uma *filha* com síndromes de Hunter e Turner, e um filho com síndrome de Hunter. Na família B, dois indivíduos de fenótipo normal tiveram duas filhas de fenótipo normal e um *filho* com síndromes de Hunter e Klinefelter. Na família C, dois indivíduos de fenótipo normal tiveram uma filha de fenótipo normal, uma *filha* com síndrome de Hunter e um filho com síndrome de Hunter. Explique, em cada família, a origem da criança indicada em *italico*.
- 6.13 Embora homens XYY tenham fenótipo normal, seria esperado que tivessem mais filhos com anormalidades dos cromossomos sexuais que homens XY? Explique.
- 6.14 Em um cromossomo da glândula salivar de *Drosophila*, a sequência das bandas é 1 2 3 4 5 6 7 8. O cromossomo homólogo com o qual essa sequência faz sinapse tem sequência 1 2 3 6 5 4 7 8. Que tipo de alteração cromossômica ocorreu? Desenhe os cromossomos em sinapse.
- 6.15 Outros cromossomos têm as seguintes sequências: (a) 1 2 5 6 7 8; (b) 1 2 3 4 4 5 6 7 8; (c) 1 2 3 4 5 8 7 6. Que tipo de alteração cromossômica ocorreu em cada um? Ilustre o pareamento desses cromossomos com um cromossomo cuja sequência é 1 2 3 4 5 6 7 8.
- 6.16 Em vegetais, os organismos heterozigotos para translocação apresentam aborto de pólen de cerca de 50%. Por quê?
- 6.17 Um cromossomo de um vegetal tem a sequência A B C D E F e outro tem a sequência M N O P Q R. Uma translocação recíproca entre esses cromossomos produziu o seguinte arranjo: A B C P Q R em um cromossomo e M N O D E F no outro. Ilustre como seria o pareamento desses cromossomos translocados com os correspondentes normais durante a meiose em indivíduo heterozigoto.
- 6.18 Em *Drosophila*, os genes *bw* e *st* estão localizados nos cromossomos 2 e 3, respectivamente. Moscas homozigotas para mutações *bw* têm olhos castanhos, moscas homozigotas para mutações *st* têm olhos escarlate e moscas homozigotas para mutações *bw* e *st* têm olhos brancos. Machos duplamente heterozigotos foram cruzados individualmente com fêmeas homozigotas *bw*, *st*. Todos os cruzamentos, exceto um, produziram quatro classes de prole: tipo selvagem, olhos castanhos, olhos escarlate e olhos brancos. A única exceção produziu apenas prole de tipo selvagem e olhos brancos. Explique a natureza dessa exceção.
- 6.19 Um menino de fenótipo normal tem 45 cromossomos, mas sua irmã, que tem síndrome de Down, tem 46. Sugira uma explicação para esse paradoxo.
- 6.20 Qual é a diferença entre um cromossomo composto e uma translocação robertsoniana?
- 6.21 Uma fêmea de *Drosophila* de corpo amarelo com cromossomos X ligados foi cruzada com um macho de olhos brancos. Os fenótipos de ambos os pais são causados por mutações recessivas ligadas ao X. Preveja os fenótipos da prole.
- 6.22 Um homem tem cromossomos 21 ligados. Sua mulher é citologicamente normal, qual é a chance de que o primeiro filho do casal tenha síndrome de Down?
- 6.23 A análise dos cromossomos politênicos de três populações de *Drosophila* mostrou três diferentes sequências de bandeamento em uma região do segundo cromossomo:

População	Sequência de bandeamento
P1	1 2 3 4 5 6 7 8 9 10
P2	1 2 3 9 8 7 6 5 4 10
P3	1 2 3 9 8 5 6 7 4 10

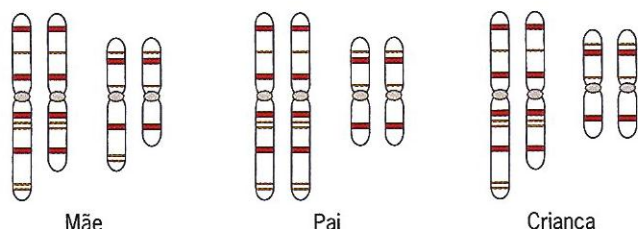
Explique as relações evolutivas entre essas populações.

- 6.24 Cada uma de seis populações de *Drosophila* em diferentes regiões geográficas apresentava um arranjo de bandas específico em um dos grandes autosomos:

- 12345678
- 12263478
- 15432678
- 14322678
- 16223478
- 154322678

Suponha que o arranjo (a) seja o original. Em que ordem provavelmente surgiram os outros arranjos? Que tipo de aberração cromossômica é responsável por cada alteração?

- 6.25 O diagrama a seguir mostra dois pares de cromossomos nos cariótipos de um homem, uma mulher e seu filho. O homem e a mulher têm fenótipo normal, mas a criança (um menino) tem uma síndrome de anormalidades, entre elas controle motor deficiente e comprometimento mental grave. Qual é a base genética do fenótipo anormal da criança? A criança é hiperploide ou hipoploide para um segmento de um de seus cromossomos?



- 6.26 Um camundongo de sexo masculino heterozigoto para uma translocação recíproca entre o cromossomo X e um autossomo é cruzado com uma fêmea de cariótipo normal. O autossomo que sofreu translocação tem um gene responsável pela coloração da pelagem. O alelo no autossomo translocado do macho é de tipo selvagem, e o alelo no autossomo não translocado é mutante; no entanto, como o alelo selvagem é dominante em relação ao mutante, a pelagem do macho é de tipo selvagem (cor escura).

A fêmea tem pelagem clara porque é homozigota para o alelo mutante do gene determinante da cor. Quando a prole do cruzamento é examinada, todos os machos têm pelagem clara e todas as fêmeas têm áreas de pelagem clara e escura. Explique esses resultados peculiares.

- 6.27 Em *Drosophila*, os genes autossômicos *cinnabar* (*cn*) e *castanho* (*bw*) controlam a produção dos pigmentos castanho e vermelho dos olhos, respectivamente. Moscas homozigotas para mutações *cinnabar* têm olhos vermelho brilhantes, moscas homozigotas para mutações *brown* têm olhos castanhos e moscas homozigotas para mutações nesses dois genes têm olhos brancos. Um macho homozigoto para mutações nos genes *cn* e *bw* tem olhos vermelho brilhantes porque uma pequena duplicação que tem o alelo selvagem de *bw* (*bw*⁺) é fixada ao cromossomo Y. Se esse macho for cruzado com uma fêmea de cariótipo normal homozigota para as mutações *cn* e *bw*, que tipos de prole serão produzidos?

- 6.28 Em *Drosophila*, as mutações de asas vestigiais (*vg*), corpo piloso (*h*) e ausência de olhos (*ey*) são recessivas nos cromossomos 2, 3, e 4, respectivamente. Os machos de tipo selvagem tratados com raios X foram cruzados com fêmeas triplamente homozigotas recessivas. Os machos da F₁ (todos com fenótipo selvagem) foram submetidos a cruzamento-teste com fêmeas homozigotas recessivas. A maioria dos machos da F₁ produziu oito classes de prole em proporções aproximadamente iguais, como seria esperado se houvesse distribuição independente dos genes *vg*, *h* e *ey*. No entanto, um macho da F₁ só produziu quatro classes de prole, cada uma delas representando cerca de um quarto do total: (1) tipo selvagem, (2) sem olhos, (3) asas vestigiais, corpo piloso e (4) asas vestigiais, corpo piloso, sem olhos. Que tipo de aberração cromossômica tem o macho excepcional da F₁ e que cromossomos foram implicados?

- 6.29 O exame citológico dos cromossomos sexuais de um homem mostrou que ele tem uma translocação para inserção. Um pequeno segmento foi deletado do cromossomo Y e inserido no braço curto do cromossomo X; esse segmento contém o gene responsável pela diferenciação do macho (*SRY*). Se esse homem casar com uma mulher de cariótipo normal, que tipos de prole terá o casal?

Genômica na Web em <http://www.ncbi.nlm.nih.gov>

1. Muitas plantas agrícolas são poliploides. Qual foi o progresso no sequenciamento dos genomas poliploides da soja (*Glycine max*), trigo (*Triticum aestivum*) e batata (*Solanum tuberosum*)?

Dica: No site, clique em *Genomes and Maps*, depois em *Genome Project* e, por fim, em *Plant Genomes*. Encontre cada espécie e leia sobre as tentativas permanentes de sequenciamento do DNA.

2. Quando triplicado, o cromossomo 21, o menor dos autossomos no genoma humano, causa a síndrome de Down. Quantos pares de nucleotídeos estão presentes nesse cromossomo? Quantos genes ele contém?

Dica: Use o *Map Viewer* para encontrar o cromossomo 21 e determinar seu tamanho e conteúdo gênico.

3. O gene da proteína precursora de amiloide, APP, está localizado no cromossomo humano 21. Essa proteína parece ter um papel importante na etiologia da doença de Alzheimer. Localize o gene *APP* no idiograma do cromossomo 21 humano. Em que banda ele está localizado?

Dica: Pesquise o APP usando a função “Find in This View”. Clique no gene destacado para descobrir mais informações sobre ele.

4. O cromossomo 21, assim como alguns outros cromossomos no genoma humano, tem constrições secundárias e uma constrição primária, situada no centrômero. O material distal à constrição secundária – isto é, afastado do centrômero em direção à extremidade mais próxima do cromossomo – é denominado satélite. Encontre a constrição secundária e o satélite no ideograma do cromossomo 21.

5. As constrições secundárias em alguns cromossomos contêm genes para o RNA ribossômico. Isso ocorre no cromossomo 21 humano?

Dica: Use a função *Map Viewer* para examinar o ideograma do cromossomo 21. Pesquise os genes de RNA ribossômicos usando a função “Find in This View”.

Ligação, Crossing over e Mapeamento Cromossômico em Eucariotos

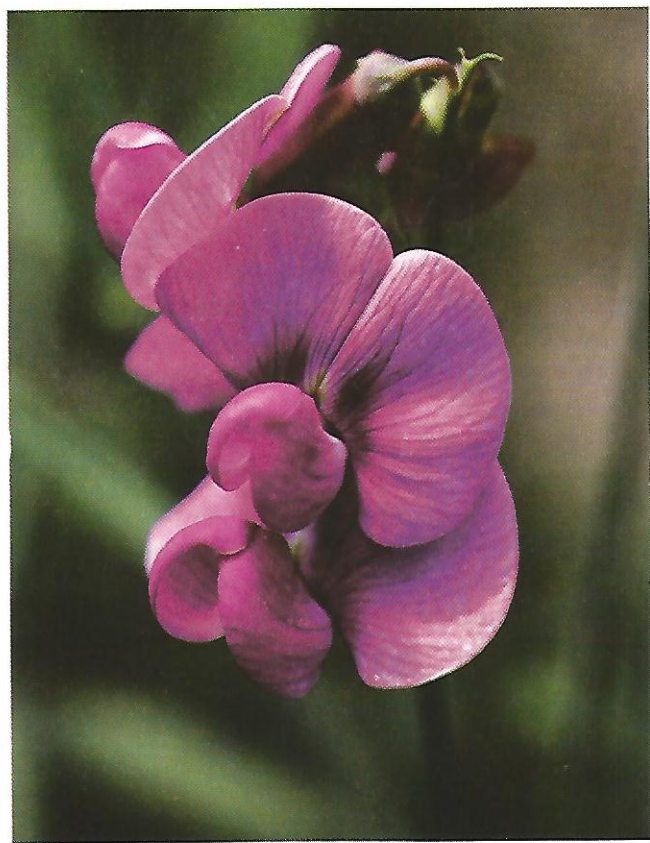
PANORAMA

- ▶ Ligação, recombinação e crossing over
- ▶ Mapeamento cromossômico
- ▶ Mapeamento citogenético
- ▶ Análise de ligação em seres humanos
- ▶ Recombinação e evolução

Primeiro mapa cromossômico do mundo

A imagem moderna da organização dos cromossomos surgiu de uma combinação de estudos genéticos e citológicos. T. H. Morgan estabeleceu os alicerces para esses estudos quando mostrou que o gene para olhos *brancos* em *Drosophila* estava localizado no cromossomo X. Logo depois, os alunos de Morgan mostraram que outros genes estavam ligados ao X e, por fim, conseguiram localizar cada um desses genes em um mapa do cromossomo. Esse mapa era uma linha reta, na qual cada gene estava situado em determinado ponto, ou *locus* (Figura 7.1). Portanto, a estrutura do mapa indicava que um cromossomo era um simples arranjo linear de genes.

O método de mapeamento dos cromossomos foi inventado por Alfred H. Sturtevant, aluno de graduação que trabalhava no laboratório de Morgan. Uma noite, em 1911, Sturtevant pôs de lado o dever de álgebra para avaliar alguns dados experimentais. Antes que o sol nascesse no dia seguinte, ele havia construído o primeiro mapa cromossômico do mundo. Como Sturtevant conseguiu determinar as localizações de cada gene no mapa? Nenhum microscópio era suficientemente eficiente para ver os genes, nem havia aparelho preciso o bastante para medir as distâncias entre eles. Na verdade, Sturtevant não usou instrumentos sofisticados em seu trabalho, apenas empregou a análise de dados de cruzamentos experimentais com *Drosophila*. Seu método era simples e refinado, e ele explorou um fenômeno que ocorre com regularidade durante a meiose. Essa metodologia criou as condições para todas as tentativas subsequentes de estudar a organização de genes em cromossomos.



A ligação entre genes foi descoberta pela primeira vez em experimentos com ervilhas-de-cheiro.

Ligação, recombinação e crossing over

Os genes que estão no mesmo cromossomo seguem juntos na meiose; no entanto, os alelos de genes ligados podem ser recombinados por *crossing over*.

Sturtevant baseou seu método de mapeamento no princípio de que genes situados no mesmo cromossomo devem ser herdados juntos. Como esses genes estão fisicamente ligados à mesma estrutura, devem seguir unidos durante

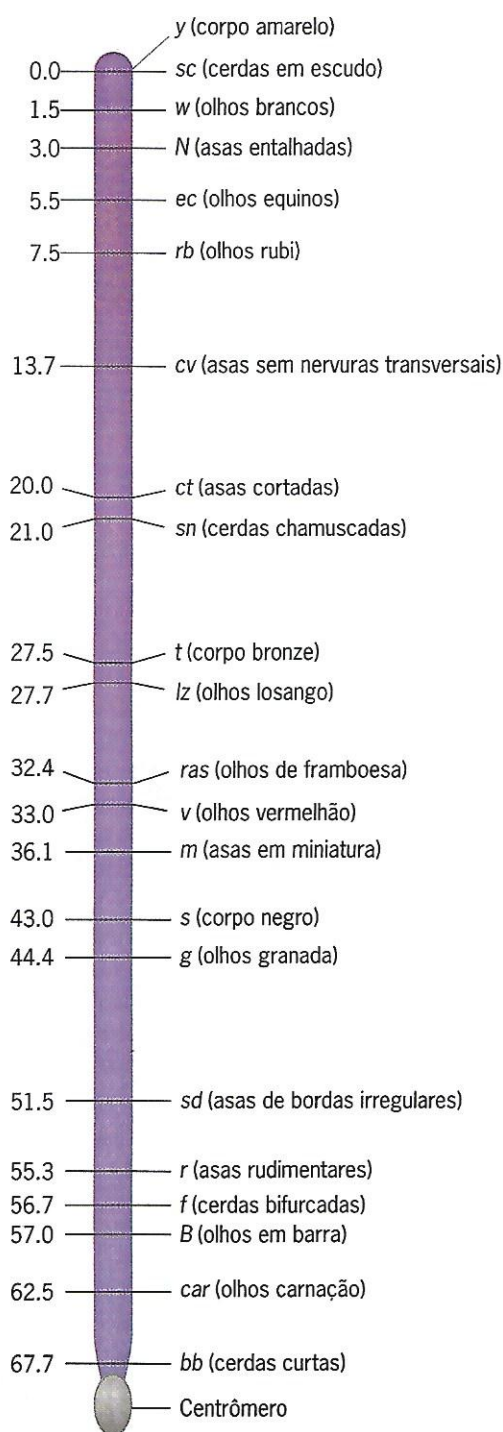
a meiose. Esse fenômeno é denominado **ligação**. Os primeiros geneticistas não tinham certeza sobre a natureza da ligação, mas alguns deles, inclusive Morgan e seus alunos, acreditavam que os genes estavam ligados uns aos outros como as contas de um cordão. Assim, é lógico que esses pesquisadores tinham em mente um modelo linear de organização dos cromossomos.

Os primeiros geneticistas também sabiam que a ligação não era absoluta. Os dados experimentais mostravam que genes no mesmo cromossomo poderiam ser separados durante a meiose e que novas combinações de genes poderiam surgir. Contudo, era difícil explicar esse fenômeno, denominado **recombinação**, pela teoria genética simples.

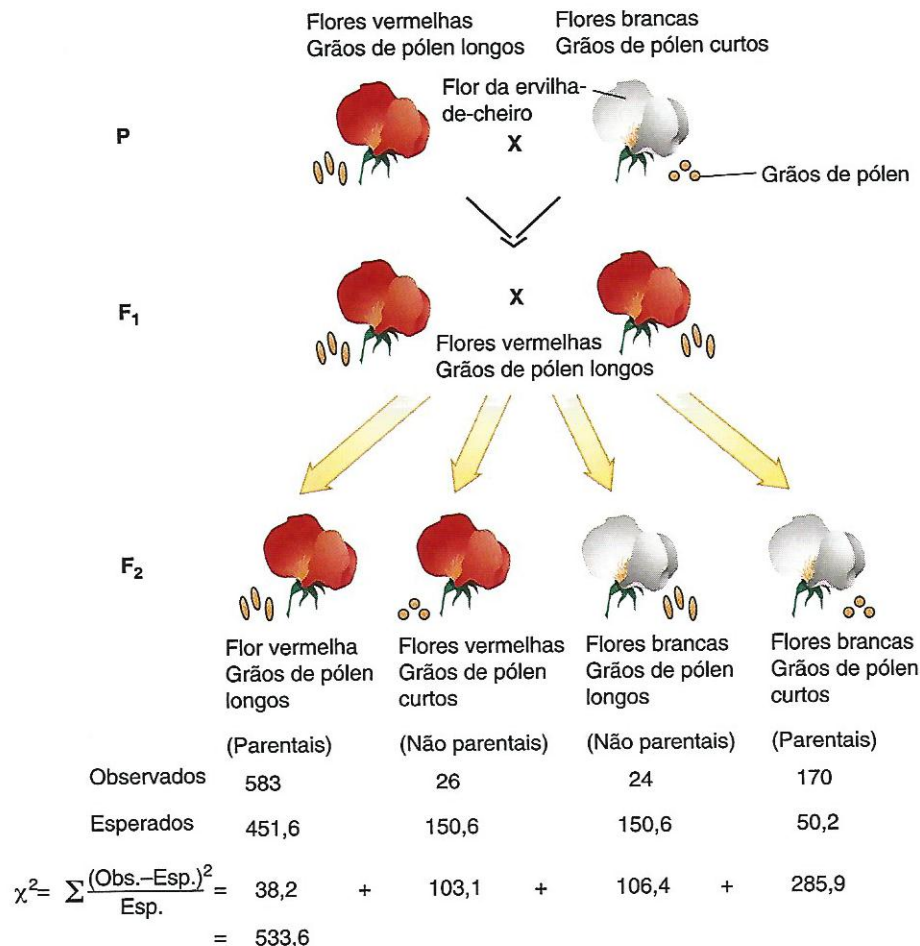
Uma hipótese afirmava que, durante a meiose, quando havia pareamento dos cromossomos homólogos, uma troca física de material separava e recombinava os genes. Essa ideia foi inspirada pela observação citológica de que era possível ver os cromossomos em configurações de pareamento que sugeriam a permuta de fragmentos entre eles. Nos pontos de permuta, os dois homólogos cruzavam-se como se cada um deles tivesse sido quebrado e, depois, fixado ao outro. Um ponto de cruzamento (*crossing over point*) foi denominado **quiasma**, palavra derivada do grego que significa “cruz”. Os geneticistas começaram a usar o termo *crossing over* para descrever o processo que criava os quiasmas – isto é, o processo real de troca entre cromossomos pareados. Eles concluíram que a recombinação – a separação de genes ligados e a formação de novas combinações gênicas – era consequência do processo físico de *crossing over*.

EVIDÊNCIAS INICIAIS DA LIGAÇÃO E RECOMBINAÇÃO

Algumas das primeiras evidências da ligação vieram de experimentos realizados por W. Bateson e R. C. Punnett (Figura 7.2). Esses pesquisadores cruzaram variedades de ervilhas-de-cheiro com duas características diferentes, cor das flores e comprimento do pólen. Plantas de flores vermelhas e grãos de pólen longos foram cruzadas com plantas de flores brancas e grãos de pólen curtos. Todas as plantas da F_1 tiveram flores vermelhas e grãos de pólen longos, indicando que os alelos para esses dois fenótipos eram dominantes. Quando as plantas da F_1 foram autofertilizadas, Bateson e Punnett observaram uma distribuição peculiar de fenótipos na prole. Em vez da proporção 9:3:3:1 esperada para dois genes de distribuição independente, obtiveram uma proporção de 24,3:1,1:1:7,1. Podemos ver o grau de discordância entre os resultados observados e os resultados esperados na parte inferior da Figura 7.2. Entre as 803 plantas da F_2 examinadas, as classes semelhantes aos pais originais (classes parentais) são super-representadas e as duas outras classes (não parentais) são sub-representadas. Em vista dessas discrepâncias



■ FIGURA 7.1 Mapa de genes no cromossomo X de *Drosophila melanogaster*.



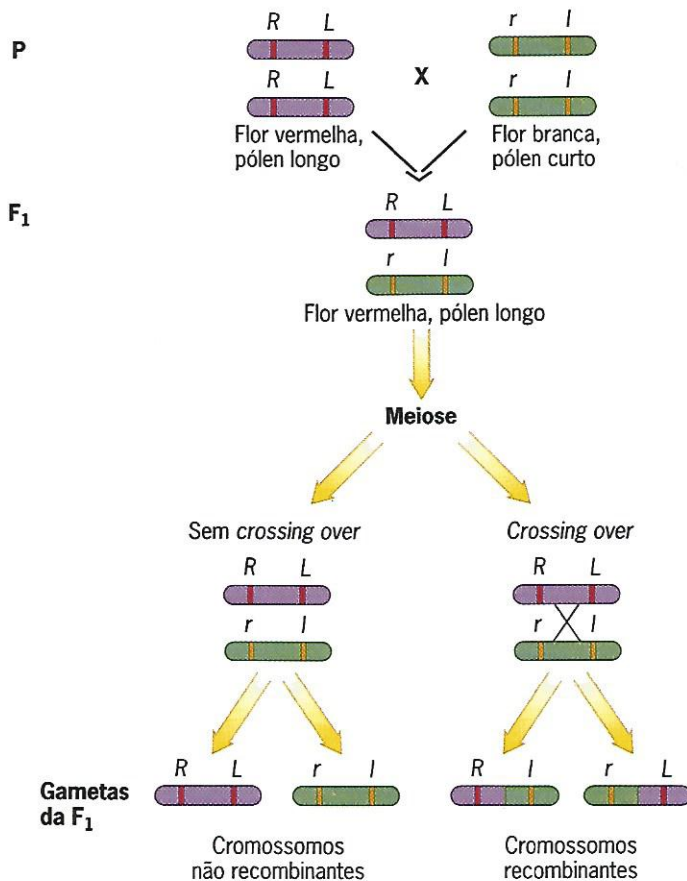
■ FIGURA 7.2 Experimento de Bateson e Punnett com ervilhas-de-cheiro. Os resultados na F₂ indicam que não há distribuição independente dos genes para cor das flores e comprimento do grão de pólen.

óbvias, é quase desnecessário calcular o qui-quadrado para testar a hipótese de distribuição independente das duas características, cor das flores e comprimento do grão de pólen. Evidentemente, eles não fizeram isso. Todavia, incluímos o cálculo do qui-quadrado na Figura 7.2 apenas para mostrar o quanto os resultados observados discordam dos resultados esperados. O valor do qui-quadrado é enorme – muito acima de 7,8, que é o valor crítico para uma distribuição de qui-quadrado com três graus de liberdade (ver Tabela 3.2). Consequentemente, precisamos rejeitar a hipótese de distribuição independente dos genes que determinam a cor das flores e o comprimento do grão de pólen.

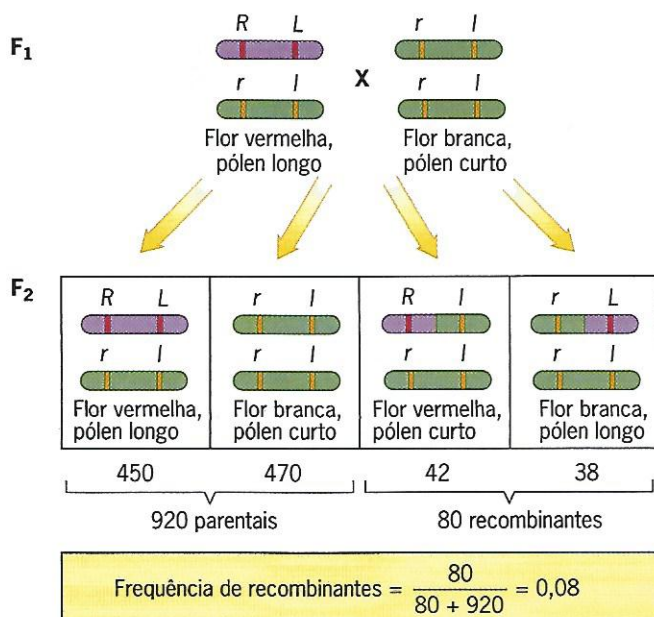
Bateson e Punnett apresentaram uma explicação complexa para seus resultados, mas que se mostrou errada. A explicação correta para a ausência de distribuição independente nos dados é que os genes para cor das flores e comprimento do pólen estão localizados no mesmo cromossomo, isto é, estão ligados. Essa explicação é apresentada no diagrama da Figura 7.3. Os alelos dos genes da cor das flores são *R* (vermelha) e *r* (branca), e os alelos do gene do comprimento do pólen são *L* (longo) e *l* (curto); os alelos *R* e *L* são dominantes. (Observe que, por motivos históricos, os símbolos dos alelos são derivados dos fenótipos dominantes, e não dos recessivos.) Como os

genes que determinam a cor das flores e o comprimento do pólen estão ligados, esperamos que as plantas da F₁ produzam dois tipos de gametas, *RL* e *rl*. No entanto, às vezes haverá um *crossing over* entre os dois genes e seus alelos serão recombinados, produzindo dois outros tipos de gametas, *Rl* e *rL*. É claro que a frequência desses dois tipos de gametas recombinantes depende da frequência de *crossing over* entre os dois genes.

Bateson e Punnett poderiam ter proposto essa explicação se tivessem realizado um cruzamento-teste em vez de um intercruzamento na F₁. Com um cruzamento-teste, a prole revelaria diretamente os tipos de gametas produzidos por plantas da F₁ duplamente heterozigotas. A Figura 7.4 apresenta a análise desse cruzamento-teste. Ervilhas-de-cheiro da F₁ duplamente heterozigotas foram cruzadas com plantas homozigotas para os alelos recessivos de ambos os genes. Entre os 1.000 organismos da prole, 920 assemelham-se a uma das linhagens parentais e os outros 80 são recombinantes. Portanto, a frequência da prole recombinante produzida pelas plantas heterozigotas da F₁ é de 80/1.000 = 0,08. Como esse é um cruzamento-teste, 0,08 também é a frequência de gametas recombinantes produzidos pelas plantas heterozigotas da F₁. Podemos usar essa frequência, geralmente denominada *frequência de recombinação*, para medir a intensidade



■ FIGURA 7.3 Hipótese de ligação entre os genes para cor das flores e comprimento do pólen em ervilhas-de-cheiro. Nas plantas da F_1 , os dois alelos dominantes, R e L , dos genes estão situados no mesmo cromossomo; seus alelos recessivos, r e l , estão situados no cromossomo homólogo.



■ FIGURA 7.4 Cruzamento-teste da ligação entre genes na ervilha-de-cheiro. Como a prole recombinante na F_2 representa 8% do total, há ligação estreita entre os genes que determinam a cor das flores e o comprimento do pólen.

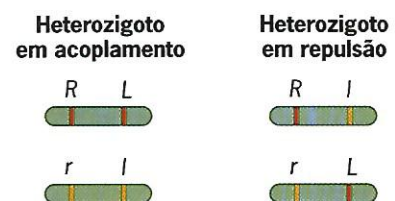
de de ligação entre genes. Os genes com ligação estreita raramente se recombinam, enquanto os genes com ligação frouxa recombinam-se com frequência. Aqui a frequência de recombinação é bastante baixa. Isso significa que o *crossing over* entre os dois genes é muito raro.

A frequência de recombinação de dois genes quaisquer nunca é maior que 50%. Esse limite máximo é alcançado quando os genes estão em cromossomos diferentes; a recombinação de 50% é, na verdade, o que ocorre quando dizemos que há distribuição independente dos genes. Por exemplo, suponhamos que os genes A e B estejam em cromossomos diferentes e que um indivíduo $AA\ BB$ seja cruzado com um indivíduo $aa\ bb$. Depois, faz-se o cruzamento-teste da prole $Aa\ Bb$ desse cruzamento com o genitor recessivo duplo. Em razão da distribuição independente dos genes A e B , a F_2 será composta de duas classes ($Aa\ Bb$ e $aa\ bb$), cujos fenótipos são iguais aos dos pais no cruzamento original, e de duas classes ($Aa\ bb$ e $aa\ Bb$), com fenótipo recombinante. Além disso, cada classe da F_2 ocorrerá com uma frequência de 25% (ver Figura 5.7). Assim, a frequência total de prole recombinante de um cruzamento-teste de dois genes em cromossomos diferentes será de 50%. Uma frequência de recombinação inferior a 50% significa que os genes estão ligados no mesmo cromossomo.

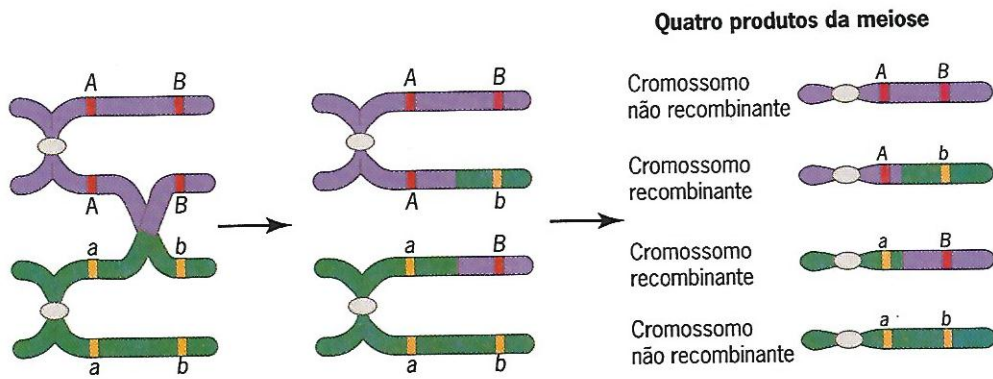
Os cruzamentos de genes ligados geralmente são apresentados em diagrama para mostrar a *fase de ligação* – a disposição dos alelos em indivíduos heterozigotos (Figura 7.5). No experimento de Bateson e Punnett com ervilhas-de-cheiro, as plantas heterozigotas da F_1 receberam dois alelos dominantes, R e L , de um genitor e dois alelos recessivos, r e l , do outro. Assim, escrevemos o genótipo dessas plantas $RL/r\ l$, em que a barra (/) separa os alelos herdados de cada genitor. Outra maneira de interpretar essa simbolização é dizer que os alelos à esquerda e à direita da barra entraram no genótipo em diferentes cromossomos homólogos, um de cada genitor. Sempre que todos os alelos dominantes estão de um lado da barra, como nesse exemplo, o genótipo tem a fase de ligação de *acoplamento*. Quando os alelos dominantes e recessivos estão divididos nos dois lados da barra, como em $R\ l/r\ L$, o genótipo tem a fase de ligação de *repulsão*. Esses termos possibilitam distinguir os dois tipos de heterozigotos duplos.

CROSSING OVER COMO BASE FÍSICA DA RECOMBINAÇÃO

Os gametas recombinantes são produzidos em consequência do *crossing over* entre cromossomos homólogos.



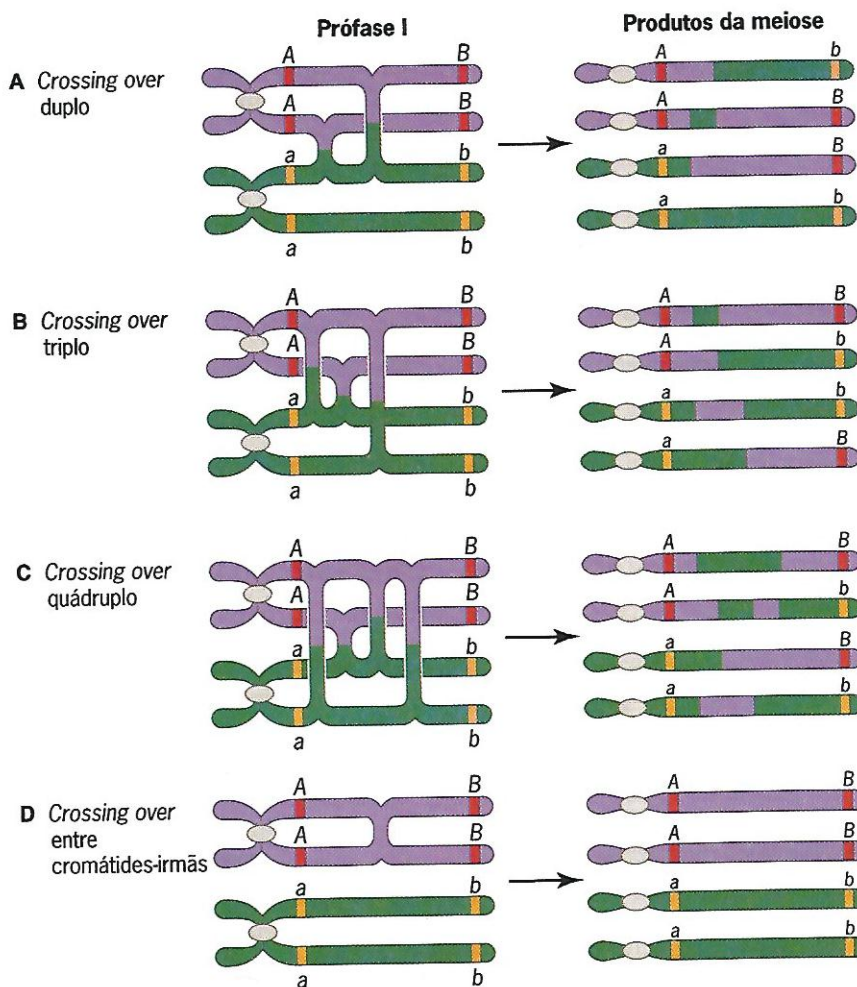
■ FIGURA 7.5 Fases de ligação de acoplamento e repulsão em heterozigotos duplos.



■ **FIGURA 7.6** *Crossing over* como base da recombinação entre genes. A permuta entre cromossomos pareados durante a meiose produz cromossomos recombinantes no fim da meiose.

Nesse processo há permuta física entre os cromossomos, como mostra a **Figura 7.6**. A permuta ocorre durante a prófase da primeira divisão meiótica, depois do pareamento dos cromossomos duplicados. Embora haja quatro cromátides homólogas, formando uma **tétrade**, apenas duas participam do *crossing over* em determinado ponto. Cada uma dessas cromátides é quebrada no local do *crossing over*, e os fragmentos se fixam novamente, produzindo os recombinantes. As outras duas cromátides não são recombinantes nesse local. Portanto, cada *crossing over* produz duas cromátides recombinantes de um total de quatro.

Observe que apenas duas cromátides participam da permuta em determinado ponto. Mas pode haver *crossing over* das outras duas cromátides em outro ponto. Assim, existe a possibilidade de múltiplas permutas em uma tétrade de cromátides (**Figura 7.7**). Por exemplo, pode haver duas, três ou até quatro permutas separadas, habitualmente denominadas *crossing overs* duplo, triplo ou quádruplo. (Iremos comentar o significado genético dessas permutas adiante neste capítulo.) Observe, porém, que uma permuta entre cromátides-irmãs não produz recombinantes genéticos porque as cromátides-irmãs são idênticas.



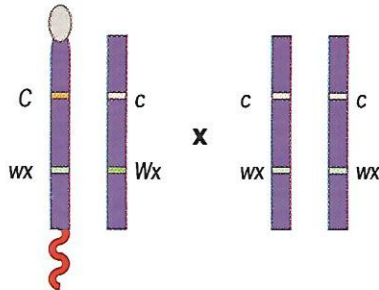
■ **FIGURA 7.7** Consequências de múltiplas permutas entre cromossomos e da permuta entre cromátides-irmãs durante a prófase I da meiose.

Qual é o responsável pela quebra de cromátides durante o *crossing over*? As quebras são causadas por enzimas que atuam no DNA das cromátides. As enzimas também são responsáveis pelo reparo dessas quebras, isto é, pela refixação dos fragmentos à outra cromátide. Apresentaremos os detalhes moleculares desse processo no Capítulo 13.

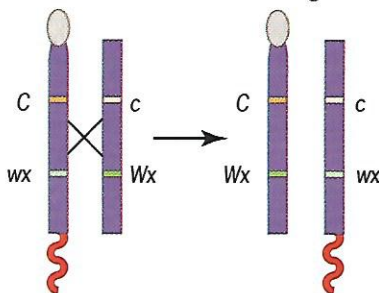
EVIDÊNCIAS DE QUE O CROSSING OVER CAUSA RECOMBINAÇÃO

Em 1931, Harriet Creighton e Barbara McClintock obtiveram evidências de que a recombinação genética estava associada à troca de material entre cromossomos. Creighton e McClintock estudaram cromossomos homólogos morfologicamente distinguíveis em milho. O objetivo era determinar se havia correlação entre a permuta física entre esses homólogos e a recombinação entre alguns dos genes presentes neles.

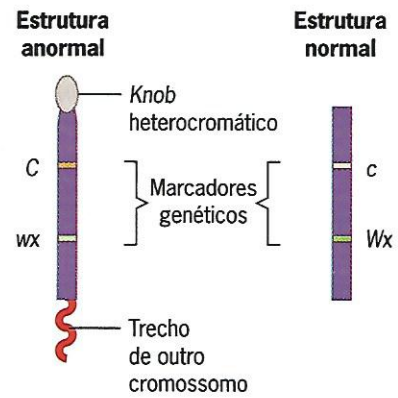
Havia duas formas de cromossomo 9 disponíveis para análise; uma era normal, e a outra tinha aberrações citológicas em cada extremidade – um *knob* heterocromático em uma extremidade e um trecho de um cromossomo diferente na outra (Figura 7.8). Essas duas formas de cromossomo 9 também foram marcadas geneticamente para detectar a recombinação. Um gene marcador controlava a cor do grão (*C*, colorido; *c*, incolor), e o outro controlava a textura do grão (*Wx*, amiláceo; *wx*, céreo). Creighton e McClintock realizaram o seguinte cruzamento-teste:



Depois, examinaram a prole recombinante para pesquisar indícios de permuta entre as duas formas diferentes de cromossomo 9. Os resultados mostraram que os recombinantes *C Wx* e *c wx* tinham um cromossomo com apenas um dos marcadores citológicos anormais; o outro marcador anormal havia sido evidentemente perdido por permuta com o cromossomo 9 normal na geração anterior:



Esses achados eram um forte indício de que a causa da recombinação era uma permuta física entre cromossomos pareados.

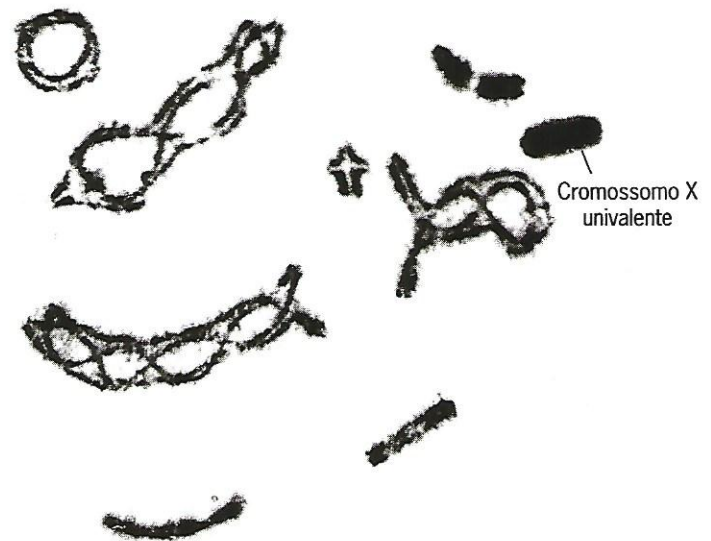


■ FIGURA 7.8 Duas formas de cromossomo 9 no milho usadas nos experimentos de Creighton e McClintock.

QUIASMAS E O MOMENTO DO CROSSING OVER

As evidências citológicas do *crossing over* são observadas no final da prófase da primeira divisão meiótica, quando é possível ver claramente os quiasmas. Nesse momento, os cromossomos pareados repelem-se ligeiramente, mantendo contato próximo apenas no centrômero e em cada quiasma (Figura 7.9). Essa separação parcial torna possível contar com precisão os quiasmas. Como seria de se esperar, os cromossomos grandes geralmente têm mais quiasmas que os cromossomos pequenos. Assim, o número de quiasmas é aproximadamente proporcional ao comprimento do cromossomo.

O surgimento de quiasmas no fim da primeira prófase meiótica poderia significar que é nesse momento que ocorre o *crossing over*. No entanto, dados de diferentes experimentos sugerem que o *crossing over* ocorre antes. Alguns desses experimentos usaram choque térmico para alterar a frequência de recombinação. Quando os



■ FIGURA 7.9 Diploteno da meiose no macho do gafanhoto *Chorthippus parallelus*. Há oito autossomos bivalentes e um cromossomo X univalente. Cada um dos quatro bivalentes menores tem um quiasma. Os demais bivalentes têm dois a cinco quiasmas.

choques térmicos foram administrados no fim da prófase, o efeito foi mínimo, mas quando administrados mais cedo, a frequência de recombinação foi modificada. Assim, o processo responsável pela recombinação, o *crossing over*, ocorre no início da prófase meiótica. Outros dados provêm de estudos moleculares sobre o momento da síntese de DNA. Embora quase todo o DNA seja sintetizado durante a intérfase que precede o início da meiose, uma pequena quantidade é produzida durante a primeira prófase meiótica. Essa síntese limitada de DNA foi interpretada como parte de um processo de reparo das cromátides fraturadas, que, como já comentamos, parece estar associado ao *crossing over*. Experimentos com controle rigoroso dos estágios mostraram que essa síntese de DNA ocorre do início ao meio da prófase, nunca depois disso. Portanto, os dados acumulados sugerem que o *crossing over* ocorre do início ao meio da prófase, muito antes que se possam ver os quiasmas.

O que, então, são os quiasmas e o que significam? A maioria dos geneticistas acredita que os quiasmas são apenas vestígios do processo real de permuta. As cromátides em que houve permuta provavelmente permanecem

entrelaçadas durante a maior parte da prófase. Por fim, esses entrelaçados se desfazem, as cromátides separam-se pelo fuso mitótico e vão para polos opostos da célula. Portanto, cada quiasma provavelmente representa um entrelaçado criado por um processo de *crossing over* mais cedo na prófase.

Mas por que existem esses entrelaçados? Muitos geneticistas acreditam que os entrelaçados criados pelo *crossing over* são uma maneira de manter unidos os membros de um bivalente durante a prófase I. A prófase I é prolongada em alguns organismos. Nas mulheres, por exemplo, pode durar até 40 anos. Sem *crossing overs*, os homólogos pareados poderiam separar-se acidentalmente durante esse longo período, o que interferiria na disjunção durante a anáfase subsequente. Erros na disjunção dos cromossomos durante a primeira divisão meiótica acabariam por produzir gametas aneuploides. Portanto, o *crossing over* parece ser um mecanismo para manter unidos os homólogos pareados de modo que, quando houver divisão, os homólogos sejam distribuídos corretamente entre as células-filhas. Assim, a possibilidade de não disjunção é reduzida ao mínimo e evita-se a aneuploidia nos gametas.

PONTOS ESSENCIAIS

- A ligação entre genes é detectada pelo desvio do esperado de acordo com o princípio de distribuição independente de Mendel
- A frequência de recombinação mede a intensidade da ligação gênica. Na ausência de ligação, essa frequência é de 50%; na ligação muito estreita, é de quase zero
- A recombinação é causada por permuta física entre cromossomos homólogos pareados no início da prófase da primeira divisão meiótica, depois da duplicação dos cromossomos
- Em um ponto qualquer ao longo de um cromossomo, só participam do processo de permuta (*crossing over*) duas das quatro cromátides de uma tétrade meiótica
- No fim da prófase I, os *crossing overs* tornam-se visíveis como quiasmas.

Mapeamento cromossômico

É possível mapear genes ligados em um cromossomo pelo estudo da frequência de recombinação de seus alelos.

O *crossing over* durante a prófase da primeira divisão meiótica tem dois resultados observáveis:

1. Formação de quiasmas no final da prófase.
2. Recombinação entre genes em lados opostos do ponto de *crossing over*.

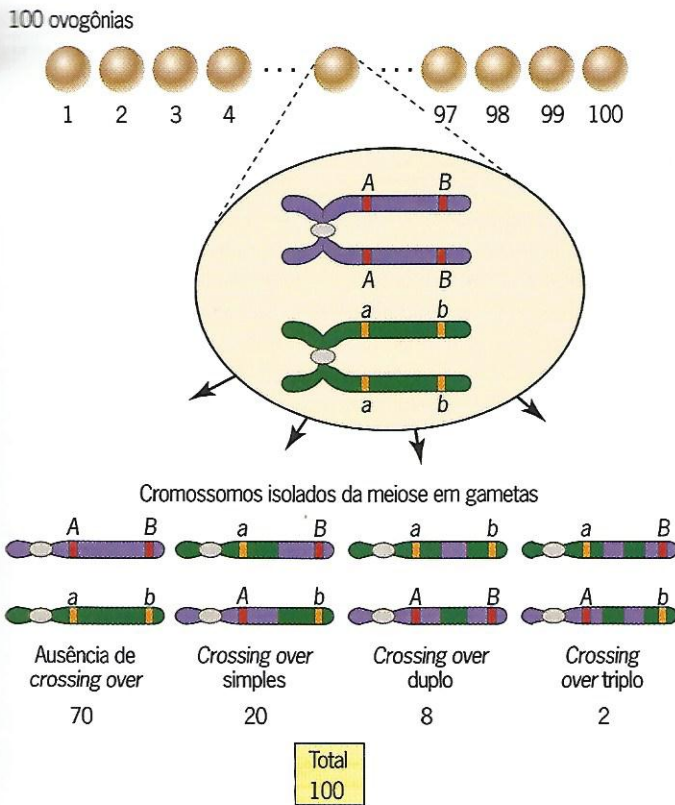
O segundo resultado, porém, só é visto na próxima geração, quando são expressos os genes nos cromossomos recombinantes.

Os geneticistas constroem mapas cromossômicos por contagem do número de *crossing overs* que ocorrem durante a meiose. No entanto, como é impossível ver o número real de *crossing overs*, não é possível contá-los diretamente. Em vez disso, é preciso estimar o número de *crossing overs* pela contagem dos quiasmas ou dos cromossomos recombinantes. Os quiasmas são contados por análise citológica,

enquanto os cromossomos recombinantes são contados por análise genética. Antes de prosseguirmos, é essencial definir o que é distância em um mapa cromossômico.

CROSSING OVER COMO MEDIDA DE DISTÂNCIA GENÉTICA

A descoberta fundamental de Sturtevant foi estimar a distância entre pontos em um cromossomo pela contagem do número de *crossing overs* entre eles. Deve haver mais *crossing overs* entre pontos distantes que entre pontos próximos. É preciso, porém, compreender o significado estatístico do número de *crossing overs*. A chance de *crossing over* entre dois pontos em determinada célula pode ser baixa, mas em uma grande população de células, é provável que esse *crossing over* ocorra várias vezes simplesmente porque há muitas oportunidades independentes. Assim, a quantidade que realmente precisamos medir é o número médio de *crossing overs* em determinada região



Número médio de crossing overs entre A e B =

$$0 \times \left(\frac{70}{100}\right) + 1 \times \left(\frac{20}{100}\right) + 2 \times \left(\frac{8}{100}\right) + 3 \times \left(\frac{2}{100}\right) = 0,42$$

■ FIGURA 7.10 Cálculo do número médio de crossing overs entre genes nos cromossomos isolados da meiose.

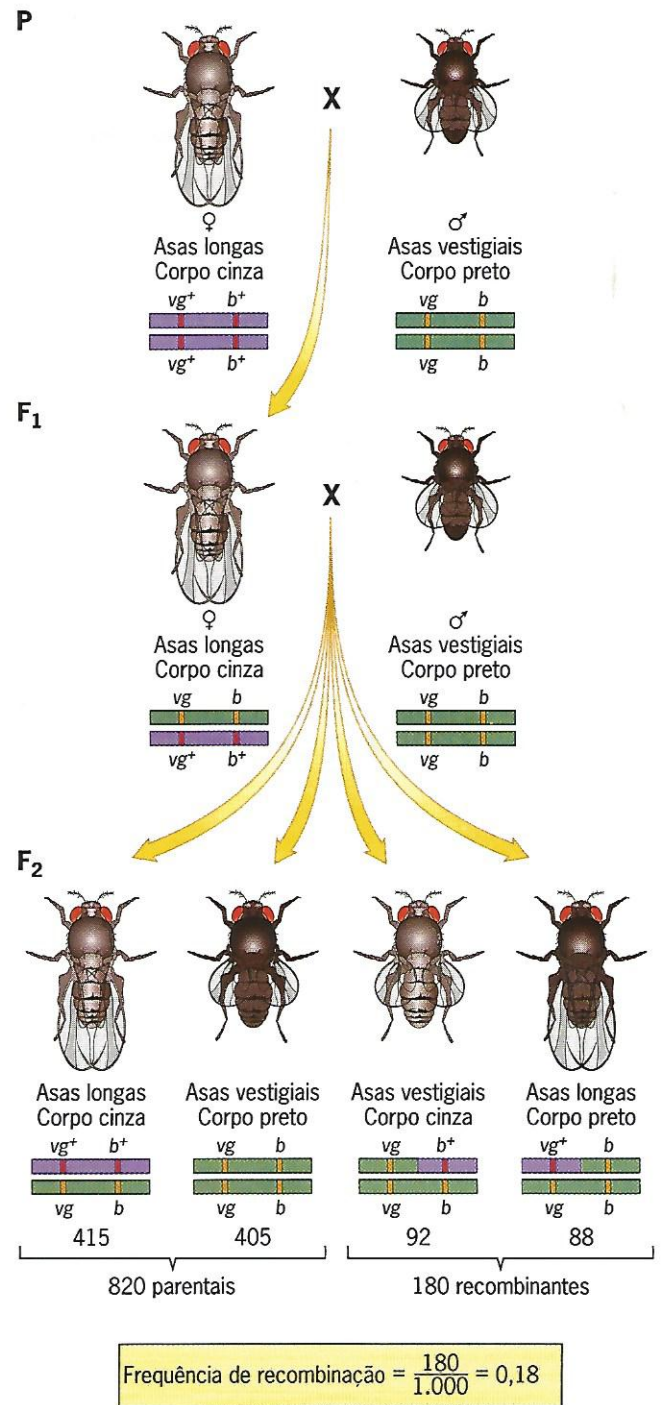
cromossômica. As distâncias no mapa genético são, na verdade, baseadas nessas médias. Essa ideia é importante o bastante para justificar uma definição formal: a distância entre dois pontos no mapa genético de um cromossomo é o número médio de crossing overs entre eles.

Uma maneira de compreendermos essa definição é considerar 100 ovogônias em meiose (Figura 7.10). Em algumas células, não haverá crossing overs entre os locais A e B; em outras, haverá um, dois ou mais crossing overs entre esses loci. No fim da meiose, haverá 100 gametas, cada um deles contendo um cromossomo com zero, um, dois ou mais crossing overs entre A e B. Estimamos a distância no mapa genético entre esses loci calculando o número médio de crossing overs nessa amostra de cromossomos. O resultado desses dados na Figura 7.10 é 0,42.

Na prática, não é possível “ver” cada ponto de permuta nos cromossomos que saem da meiose. Em vez disso, deduzimos sua existência observando a recombinação dos alelos ao seu lado. Um cromossomo no qual houve recombinação de alelos surgiu obrigatoriamente por crossing over. Portanto, a contagem de cromossomos recombinantes é uma maneira de contar os pontos de crossing over.

MAPEAMENTO DE RECOMBINAÇÃO COM UM CRUZAMENTO-TESTE DE DOIS PONTOS

Para ilustrar a técnica de mapeamento, consideremos o cruzamento-teste de dois pontos na Figura 7.11. As fê-



■ FIGURA 7.11 Experimento com dois genes ligados, *vg* (asas vestigiais) e *b* (corpo preto), em *Drosophila*.

meas de *Drosophila* de tipo selvagem foram cruzadas com machos homocigotos para duas mutações autossômicas – *vestigial* (*vg*), que produz asas curtas, e *black* (*b*), que produz corpo preto. Todas as moscas da F₁ tinham asas longas e corpos cinza; assim, os alelos selvagens (vg^+ e b^+) são dominantes. Em seguida, fez-se o cruzamento-teste das fêmeas da F₁ com machos de corpo preto e asas curtas, e a prole da F₂ foi classificada por fenótipo e contada. Como mostram os dados, há quatro classes fenotípicas, duas abundantes e duas raras. As classes abundantes tinham o mesmo fenótipo que os pais originais, e as classes raras tinham fenótipo recombinante. Sabemos que os ge-

nes *vestigial* e *black* estão ligados porque os recombinantes são muito menos de 50% da prole total. Portanto, é obrigatório que esses genes estejam no mesmo cromossomo. Para determinar a distância entre eles, é preciso estimar o número médio de *crossing overs* nos gametas das fêmeas heterozigotas duplas da F_1 . Podemos fazer isso calculando a frequência de moscas recombinantes na F_2 e observando que cada mosca desse tipo herdou um cromossomo em que houve um *crossing over* entre *vg* e *b*. Portanto, o número médio de *crossing overs* em toda a amostra da prole é

$$\begin{array}{ccc} \text{não recombinantes} & & \text{recombinantes} \\ (0) \times 0,82 & + & (1) \times 0,18 = 0,18 \end{array}$$

Nessa expressão, o número de *crossing overs* em cada classe de moscas está entre parênteses; o outro número é a frequência dessa classe. Obviamente, a prole não recombinante não acrescenta cromossomos em *crossing over* aos dados, mas nós a incluímos no cálculo para enfatizar que é preciso calcular o número médio de *crossing overs* usando todos os dados, não só os dados dos recombinantes.

Essa análise simples indica que, em média, 18 de 100 cromossomos isolados da meiose tinham um *crossing over* entre *vg* e *b*. Assim, *vg* e *b* estão separados por 18 unidades de mapa genético. Às vezes, os geneticistas chamam a unidade de mapa de **centiMorgan**, abreviado como cM, em homenagem a T. H. Morgan; 100 centiMorgans correspondem a um Morgan (M). Portanto, podemos dizer que *vg* e *b* estão distantes 18 cM (ou 0,18 M). Note que a distância no mapa é igual à frequência de recombinação, escrita na forma de porcentagem. Mais tarde, veremos que, quando se aproxima de 0,5, a frequência de recombinação subestima a distância no mapa. Teste sua compreensão dos princípios do mapeamento de recombinação acompanhando o exercício do Boxe Resolva: Mapeamento de dois genes com dados do cruzamento-teste.

MAPEAMENTO DE RECOMBINAÇÃO COM UM CRUZAMENTO-TESTE DE TRÊS PONTOS

Também podemos usar o mapeamento de recombinação com dados dos cruzamentos-teste com participação de mais de dois genes. A Figura 7.12 ilustra um experimento de C. B. Bridges e T. M. Olbrycht, que cruzaram machos de *Drosophila* de tipo selvagem com fêmeas homozigotas para três mutações recessivas ligadas ao X – cerdas *scute* (em escudo [*sc*]), olhos *echinus* (equinos [*ec*]) e asas *crossveinless* (sem nervuras transversais [*cv*]). Em seguida, inter cruzaram a prole da F_1 para produzir moscas F_2 , que classificaram e contaram. Observamos que as fêmeas da F_1 nesse inter cruzamento tinham as três mutações recessivas em um de seus cromossomos X e os alelos selvagens dessas mutações no outro cromossomo X. Além disso, os machos da F_1 tinham as três mutações recessivas no único cromossomo X. Assim, esse inter cruzamento foi equivalente a um cruzamento-teste no qual os três genes

Resolva!

Mapeamento de dois genes com dados do cruzamento-teste

No milho, o gene para cor da folha tem dois alelos, *g* recessivo para folhas verdes e *G* dominante para folhas roxas, e o gene para altura do caule tem dois alelos, *s* recessivo para caule baixo e *S* dominante para caule alto. Uma planta de folhas verdes e caule baixo foi cruzada com outra de folhas roxas e caule alto. Todas as plantas da F_1 desse cruzamento apresentaram folhas roxas e caules altos. Quando retrocruzadas com plantas de folhas verdes e caules baixos, produziram uma F_2 na qual, em um total de 200 plantas, observaram-se quatro classes fenotípicas: (1) folhas verdes, caules baixos, 75; (2) folhas roxas, caules altos, 79; (3) folhas verdes, caules altos, 24; e (4) folhas roxas, caules baixos, 22. (a) Qual é a evidência da ligação entre os genes para cor das folhas e altura do caule? (b) Qual é a fase de ligação dos alelos dominante e recessivo desses dois genes nas plantas da F_1 ? (c) Entre as plantas da F_2 , qual é a frequência de recombinação? (d) Qual é a distância em centiMorgans entre os genes da cor das folhas e altura do caule?

► Leia a resposta do problema no site <http://gen-io.grupogen.com.br>.

nas fêmeas da F_1 estavam presentes na configuração em acoplamento.

As moscas da F_2 do inter cruzamento eram de oito classes fenotípicas, duas parentais e seis recombinantes. As classes parentais eram, sem dúvida, as mais numerosas. As classes recombinantes, menos numerosas, representavam diferentes tipos de cromossomos produzidos por *crossing over*. Para identificar os *crossing overs* que produziram cada tipo de recombinante, precisamos primeiro determinar a ordem dos genes no cromossomo.

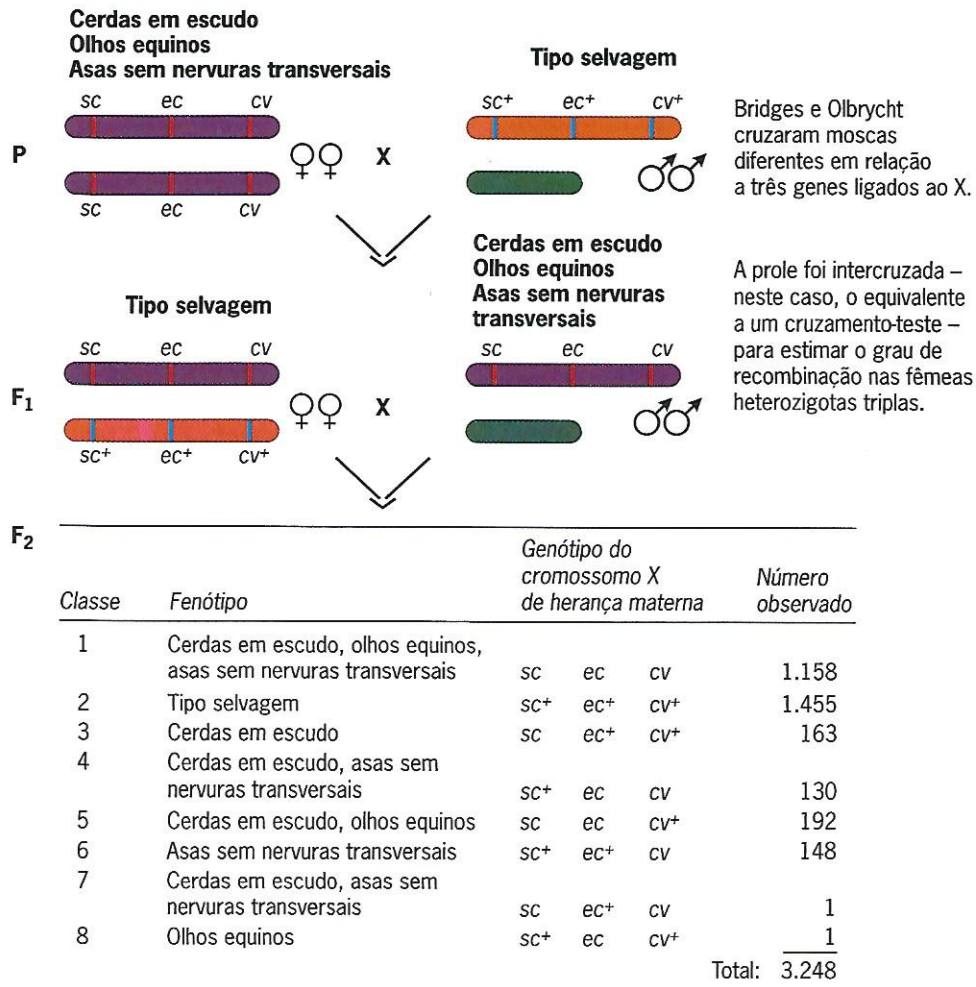
Determinação da ordem dos genes

Existem três ordens de genes possíveis:

1. *sc*–*ec*–*cv*
2. *ec*–*sc*–*cv*
3. *ec*–*cv*–*sc*

Outras possibilidades, como *cv*–*ec*–*sc*, são iguais a uma dessas porque não há distinção entre as extremidades esquerda e direita do cromossomo. Qual das ordens é a certa?

Para responder a essa pergunta, precisamos analisar meticulosamente as seis classes recombinantes. Quatro delas foram produzidas obrigatoriamente por um *crossing over* simples em uma das duas regiões delimitadas pelos genes. As outras duas foram produzidas obrigatoriamente por *crossing over* duplo – uma permuta em cada uma das duas regiões. Como um *crossing over* duplo altera o gene do meio em relação aos marcadores genéticos de cada lado dele, temos, em princípio, um método para determinar a ordem dos genes. Intuitivamente, também sabemos que um *crossing over* duplo deve ser muito menos



■ **FIGURA 7.12** Cruzamento de três pontos de Bridges e Olbrycht com os genes ligados ao X (*sc* (cerdas em escudo), *ec* (olhos equinos) e *cv* (asas sem nervuras transversais) em *Drosophila*. Dados retirados de Bridges, C. B., e Olbrycht, T. M., 1926. *Genetics* 11: 41.

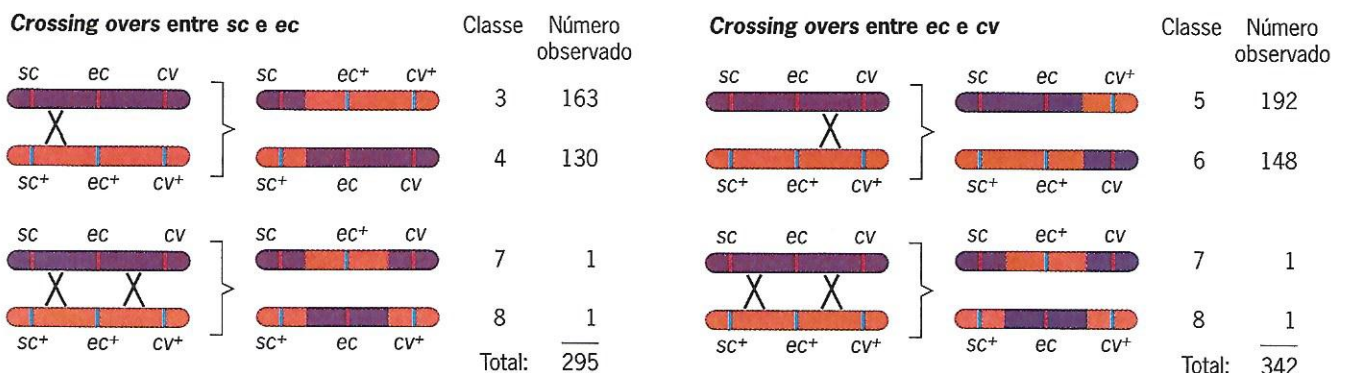
frequente que um *crossing over* simples. Consequentemente, entre as seis classes recombinantes, as duas raras têm de representar os cromossomos com *crossing over* duplo.

Em nossos dados, as classes de *crossing over* duplo, raras, são 7 (*sc ec⁺ cv*) e 8 (*sc⁺ ec cv⁺*), cada uma delas com apenas uma mosca (Figura 7.12). Comparando essas classes com as classes parentais 1 (*sc ec cv*) e 2 (*sc⁺ ec⁺ cv⁺*), verificamos que o alelo *echinus* foi trocado em relação a *scute* e *cross-veinless*. Consequentemente, o gene *echinus* tem de estar

localizado entre os outros dois. Assim, a ordem correta dos genes é (1) *sc–ec–cv*.

Cálculo das distâncias entre genes

Estabelecida a ordem dos genes, podemos determinar as distâncias entre genes adjacentes. Mais uma vez, o método é calcular o número médio de *crossing overs* em cada região cromossômica (Figura 7.13).



$$\text{Distância de mapa} = \frac{295}{3.248} = 0,091 \text{ Morgan} = 9,1 \text{ centiMorgans}$$

$$\text{Distância de mapa} = \frac{342}{3.248} = 0,105 \text{ Morgan} = 10,5 \text{ centiMorgans}$$

■ **FIGURA 7.13** Cálculo de distâncias de mapa genético a partir dos dados de Bridges e Olbrycht. A distância entre cada par de genes é obtida por estimativa do número médio de *crossing overs*.

Podemos obter o comprimento da região entre *sc* e *ec* identificando as classes recombinantes em que havia *crossing over* entre esses genes. Existem quatro dessas classes: 3 (*sc ec⁺ cv⁺*), 4 (*sc⁺ ec cv*), 7 (*sc ec⁺ cv*) e 8 (*sc⁺ ec cv⁺*). As classes 3 e 4 tiveram um *crossing over* simples entre *sc* e *ec*, e as classes 7 e 8 tiveram dois *crossing overs*, um entre *sc* e *ec* e outro entre *ec* e *cv*. Portanto, podemos usar as frequências dessas quatro classes para estimar o número médio de *crossing overs* entre *sc* e *ec*.

$$\frac{\text{Classe 3} \quad \text{Classe 4} \quad \text{Classe 7} \quad \text{Classe 8}}{163 \quad + \quad 130 \quad + \quad 1 \quad + \quad 1}{\text{Total}} = \frac{295}{3.248} = 0,091$$

Assim, a cada 100 cromossomos originados da meiose nas fêmeas da F_1 , 9,1 tinham um *crossing over* entre *sc* e *ec*. Portanto, a distância entre esses genes é de 9,1 unidades de mapa (ou, se preferir, 9,1 centiMorgans).

Da mesma maneira, podemos calcular a distância entre *ec* e *cv*. Quatro classes recombinantes apresentavam um *crossing over* nessa região: 5 (*sc ec cv⁺*), 6 (*sc⁺ ec⁺ cv*), 7 e 8. Os recombinantes duplos também estão incluídos aqui porque um dos dois *crossing overs* ocorreu entre *ec* e *cv*. A frequência total dessas quatro classes é:

$$\frac{\text{Classe 5} \quad \text{Classe 6} \quad \text{Classe 7} \quad \text{Classe 8}}{192 \quad + \quad 148 \quad + \quad 1 \quad + \quad 1}{\text{Total}} = \frac{342}{3.248} = 0,105$$

Consequentemente, *ec* e *cv* estão distantes 10,5 unidades de mapa.

Combinando os dados acerca das duas regiões, obtemos o mapa

$$sc-9,1-ec-10,5-cv$$

As distâncias de mapa calculadas dessas formas são aditivas. Assim, é possível estimar a distância entre *sc* e *cv* somando os comprimentos dos dois intervalos entre eles no mapa:

$$9,1 \text{ cM} + 10,5 \text{ cM} = 19,6 \text{ cM}$$

Também podemos fazer essa estimativa calculando diretamente o número médio de *crossing overs* entre esses genes:

Classes sem <i>crossing over</i>	Classes com <i>crossing over</i> simples	Classes com <i>crossing over</i> duplo
1 e 2	3, 4, 5 e 6	7 e 8

$$(0) \times 0,805 + (1) \times 0,195 + (2) \times 0,0006 = 0,196$$

Aqui o número de *crossing overs* é apresentado entre parênteses, e seu multiplicador é a frequência combinada das classes com esse número de *crossing overs*. Em outras palavras, cada classe recombinante contribui para a distância de mapa de acordo com o produto de sua frequência e o número de *crossing overs* que representa.

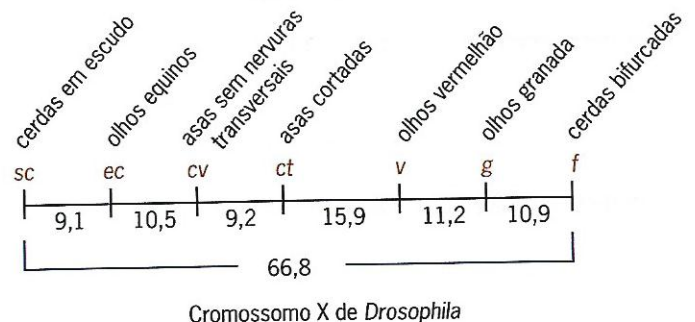
Bridges e Olbrycht estudaram sete genes ligados ao X em seu experimento de recombinação: *sc*, *ec*, *cv*, *ct* (asas cortadas [*cut*]), *v* (olhos vermelho [*vermilion*]), *g* (olhos granada [*garnet*]) e *f* (cerdas bifurcadas [*forked*]). Calculando as frequências de recombinação entre cada par de genes adjacentes, eles construíram um mapa de um grande segmento do cromossomo X (Figura 7.14); *sc* estava em uma extremidade, e *f* na outra. Todos os sete genes que Bridges e Olbrycht estudaram eram, na verdade, *marcadores* de sítios específicos no cromossomo X. Somando todos os intervalos de mapa entre esses marcadores, eles estimaram que o comprimento total do segmento mapeado era 66,8 cM. Assim, o número médio de *crossing overs* nesse segmento era 0,668.

Interferência e coeficiente de coincidência

O cruzamento de três pontos tem uma importante vantagem em relação ao cruzamento de dois pontos: torna possível detectar *crossing overs* duplos e determinar se as permutas em regiões adjacentes são independentes. Por exemplo, um *crossing over* na região entre *sc* e *ec* (região I no mapa do cromossomo X) é independente de um *crossing over* na região entre *ec* e *cv* (região II)? Ou um *crossing over* inibe a ocorrência de outro próximo dele?

Para responder essas perguntas, é preciso calcular a frequência esperada de *crossing overs* duplos, com base no conceito de independência. Podemos fazer isso multiplicando as frequências de *crossing over* para duas regiões cromossômicas adjacentes. Por exemplo, na região I do mapa de Bridges e Olbrycht, a frequência de *crossing over* foi $(163 + 130 + 1 + 1)/3.248 = 0,091$, e na região II foi $(192 + 148 + 1 + 1)/3.248 = 0,105$. Na hipótese da independência, a frequência esperada de *crossing overs* duplos no intervalo entre *sc* e *cv* seria, portanto, $0,091 \times 0,105 = 0,0095$. Agora podemos comparar essa frequência à frequência observada, que foi de $2/3.248 = 0,0006$. *Crossing overs* duplos entre *sc* e *cv* eram muito menos frequentes que o esperado. Esse resultado sugere que um *crossing over* inibia a ocorrência de outro próximo, um fenômeno denominado *interferência*. O grau de interferência é medido habitualmente pelo *coeficiente de coincidência*, *c*, que é a razão entre a frequência observada e a frequência esperada de *crossing overs* duplos:

$$c = \frac{\text{frequência observada de } \textit{crossing overs} \text{ duplos}}{\text{frequência esperada de } \textit{crossing overs} \text{ duplos}} = \frac{0,0006}{0,0095} = 0,063$$



■ FIGURA 7.14 Mapa de Bridges e Olbrycht de sete genes ligados ao X em *Drosophila*. As distâncias são apresentadas em centiMorgans.

O nível de interferência, cujo símbolo é I , é calculado por $I = 1 - c = 0,937$.

Como nesse exemplo o coeficiente de coincidência é próximo de zero, o valor mínimo possível, a interferência era muito forte (I está próximo de 1). No outro extremo, um coeficiente de coincidência igual a um implicaria ausência de interferência; isto é, significaria que os *crossing overs* foram independentes.

Muitos estudos mostraram que a interferência é forte quando as distâncias no mapa são menores que 20 cM; assim, raramente há *crossing overs* duplos em regiões cromossômicas curtas. No entanto, em regiões longas, a interferência enfraquece até o ponto em que os *crossing overs* ocorrem de maneira mais ou menos independente. Portanto, o grau de interferência é uma função da distância no mapa.

Uma vez construído o mapa genético, é possível usá-lo para prever os resultados dos experimentos. Veja como são feitas as previsões baseadas em mapa acompanhando o Problema resolvido: Uso do mapa genético para prever o resultado de um cruzamento.

FREQUÊNCIA DE RECOMBINAÇÃO E DISTÂNCIA NO MAPA GENÉTICO

Nas seções anteriores, abordamos a construção de mapas cromossômicos a partir de dados sobre a recombinação de marcadores genéticos. Esses dados tornam possível deduzir onde ocorreram os *crossing overs* em uma amostra de cromossomos. Localizando e contando esses *crossing overs*,

é possível estimar as distâncias entre os genes e representá-los em um mapa cromossômico.

Esse método é eficaz desde que os genes estejam bem próximos. Quando estão afastados, porém, a frequência de recombinação pode não refletir a verdadeira distância no mapa (Figura 7.15). Consideremos, por exemplo, os genes nas extremidades do mapa do cromossomo X feito por Bridges e Olbrycht; *sc*, na extremidade esquerda, estava 66,8 cM distante de *f*, na extremidade direita. No entanto, a frequência de recombinação entre *sc* e *f* era de 50%, o valor máximo possível. Usando essa frequência para estimar a distância no mapa, concluiríamos que *sc* e *f* estavam distantes 50 unidades de mapa. É claro que a distância calculada pela soma dos comprimentos das regiões interpostas no mapa, 66,8 cM, é muito maior.

Esse exemplo mostra que a verdadeira distância genética, que depende do número médio de *crossing overs* em um cromossomo, pode ser muito maior que a frequência de recombinação observada. Pode haver múltiplos *crossing overs* entre genes muito afastados, e alguns deles podem não produzir cromossomos geneticamente recombinantes (Figura 7.16). Para verificar isso, vamos supor que haja um *crossing over* simples entre duas cromátides em uma tétrade, causando a recombinação dos marcadores genéticos flanqueadores. Se houver outro *crossing over* entre essas mesmas duas cromátides, os marcadores flanqueadores voltarão à configuração original; na prática, o segundo *crossing over* anula o efeito do primeiro, reconvertendo as cromátides recombinantes em não recombinantes. Assim, ainda que tenha havido dois *crossing overs*,

PROBLEMA RESOLVIDO



Uso do mapa genético para prever o resultado de um cruzamento

PROBLEMA

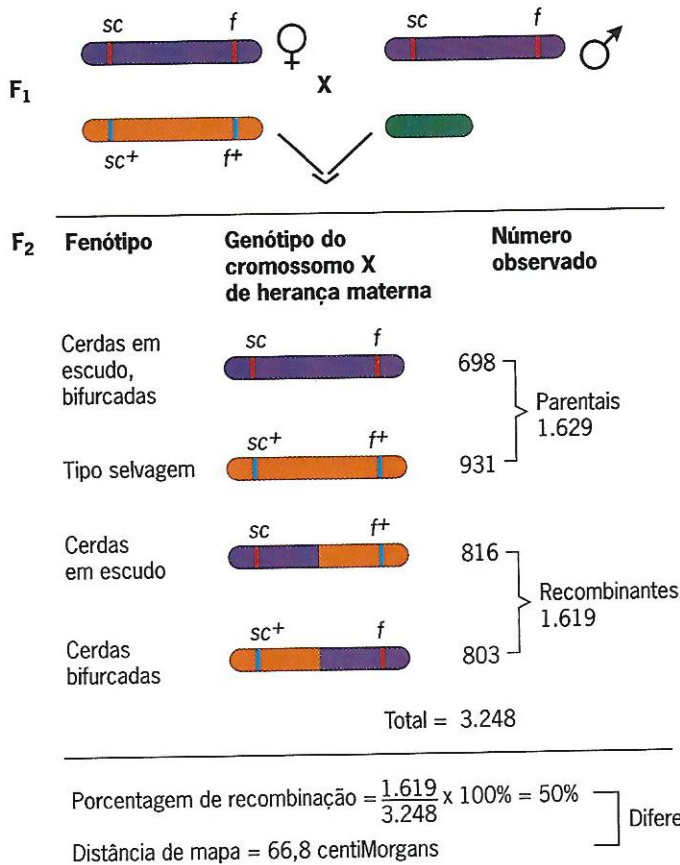
Os genes *r*, *s* e *t* estão no meio do cromossomo X de *Drosophila*; *r* está 15 cM à esquerda de *s*, e *t* está 20 cM à direita de *s*. Nessa região, o coeficiente de coincidência (c) é 0,2. Um geneticista quer criar um cromossomo X que tenha os alelos mutantes recessivos dos três genes. Um estoque é homozigoto para *r* e *t*, e o outro é homozigoto para *s*. Cruzando os dois estoques, o geneticista obtém fêmeas que são heterozigotas triplos, $r s^+ t/r^+ s t^+$. Essas fêmeas são, então, cruzadas com machos de tipo selvagem. Se o geneticista examinar 10.000 filhos machos dessas fêmeas, quantos deles serão mutantes triplos, $r s t$?

FATOS E CONCEITOS

1. Quando os intervalos no mapa são pequenos (< 20 cM), a distância no mapa é igual à frequência de um *crossing over* simples no intervalo.
2. O coeficiente de coincidência é igual à razão frequência observada/frequência esperada de *crossing overs* duplos.
3. A frequência esperada de *crossing overs* duplos é calculada com base na suposição de que os dois *crossing overs* são independentes.
4. Os machos herdam o cromossomo X das mães.

ANÁLISE E SOLUÇÃO

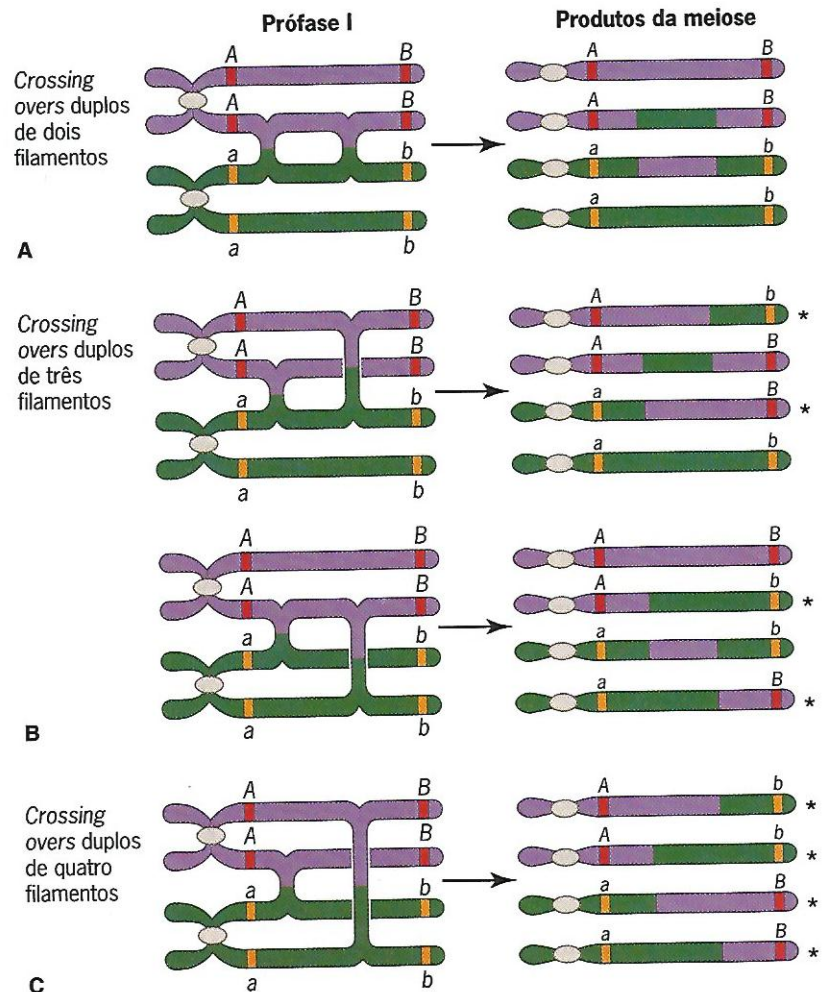
Os machos mutantes triplos só são produzidos se houver um *crossing over* duplo nas fêmeas $r s^+ t/r^+ s t^+$ cruzadas com machos de tipo selvagem. A frequência desses *crossing overs* duplos é uma função das duas distâncias no mapa (15 cM e 20 cM) e do nível de interferência, medido pelo coeficiente de coincidência (aqui $c = 0,2$). Como $c = \text{frequência observada} / \text{frequência esperada de } \textit{crossing overs} \text{ duplos}$, podemos calcular a frequência observada de *crossing overs* duplos depois de uma reorganização algébrica simples: frequência observada de *crossing overs* duplos = $c \times \text{frequência esperada de } \textit{crossing overs} \text{ duplos}$. A frequência esperada de *crossing overs* duplos é calculada a partir das distâncias no mapa, supondo-se que os *crossing overs* nos intervalos adjacentes do mapa sejam independentes: $0,15 \times 0,20 = 0,03$. Assim, em 10.000 filhos machos, $0,2 \times 3\%$ devem ter um cromossomo X que teve um *crossing over* entre os genes *r* e *s* e outro *crossing over* entre os genes *s* e *t*. No entanto, apenas metade desses 60 filhos – isto é, 30 – terão o cromossomo X mutante triplo; os outros 30 terão o tipo selvagem triplo.



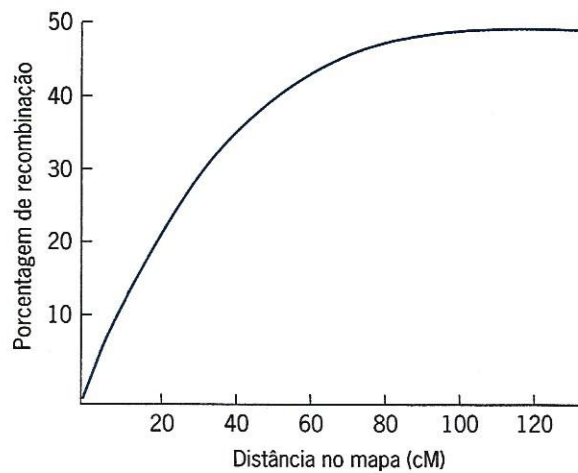
■ **FIGURA 7.15** Discrepância entre a distância de mapa e a porcentagem de recombinação. A distância de mapa entre os genes *sc* e *f* é maior que o porcentual observado de recombinação entre eles.

overs nessa tétrade, nenhuma das cromátides dela será recombinante para os marcadores flanqueadores.

Esse segundo exemplo mostra que um *crossing over* duplo não pode contribuir para a frequência de recombinação, embora contribua para o número médio de permutas em um cromossomo. Um *crossing over* quádruplo teria o mesmo efeito. Essas e outras permutas múltiplas são responsáveis pela discrepância entre a frequência de recombinação e a distância no mapa genético. Na prática, essa discrepância é pequena quando a distância é inferior a 20 cM. Nessas distâncias, a interferência é grande o suficiente para inibir quase todas as permutas múltiplas, e a frequência de recombinação é uma boa estimativa da verdadeira distância genética. Quando a distância é maior que 20 cM, essas duas quantidades divergem, principalmente porque se torna muito mais provável a ocorrência de permutas múltiplas. A Figura 7.17 mostra a relação matemática entre a frequência de recombinação e a distância no mapa genético.



■ **FIGURA 7.16** Consequências do *crossing over* duplo entre dois loci. Os cromossomos recombinantes são indicados por um asterisco. (A) Os *crossing overs* duplos de dois filamentos produzem apenas cromossomos não recombinantes. (B) *Crossing overs* duplos de três filamentos produzem metade dos cromossomos recombinantes e metade não recombinantes. (C) Os *crossing overs* duplos de quatro filamentos produzem apenas cromossomos recombinantes.



■ **FIGURA 7.17** Relação entre a frequência de recombinação e a distância no mapa genético. Abaixo de 20 cM, há uma relação aproximadamente linear entre a porcentagem de recombinação e a distância no mapa; acima de 20 cM, a porcentagem de recombinação subestima a distância no mapa.

PONTOS ESSENCIAIS

- Os mapas genéticos de cromossomos são baseados no número médio de crossing overs ocorridos durante a meiose
- As distâncias no mapa genético são estimadas pelo cálculo da frequência de recombinação entre genes em cruzamentos experimentais
- Frequências de recombinação abaixo de 20% estimam diretamente a distância no mapa; no entanto, frequências de recombinação acima de 20% subestimam a distância no mapa porque crossing overs múltiplos nem sempre produzem cromossomos recombinantes.

Mapeamento citogenético

Os geneticistas desenvolveram técnicas de localização dos genes nos mapas citológicos dos cromossomos.

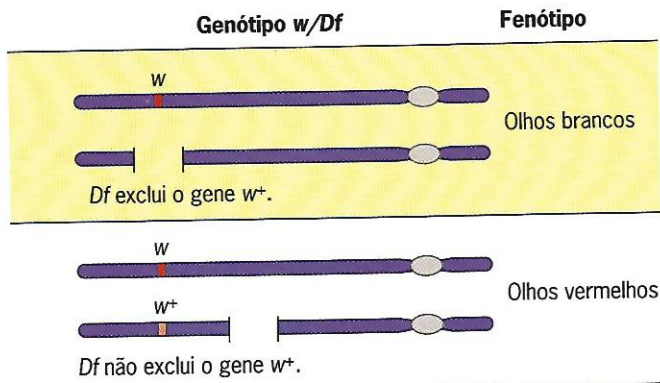
O mapeamento da recombinação torna possível identificar as posições relativas dos genes usando a frequência de *crossing over* como medida de distância. No entanto, não possibilita a localização de genes em relação aos pontos de referência citológicos, como bandas, nos cromossomos. Esse tipo de localização requer outro procedimento, com estudo dos efeitos fenotípicos de rearranjos cromossômicos, como deleções e duplicações. Como é possível reconhecer citologicamente esses tipos de rearranjos, pode-se correlacionar seus efeitos fenotípicos a determinadas regiões ao longo de um cromossomo. Se for possível associar esses efeitos fenotípicos a genes já posicionados em um mapa de recombinação, podem-se vincular as posições desses genes no mapa a locais no mapa citológico de um cromossomo. Esse processo de *mapeamento citogenético* foi totalmente desenvolvido no estudo genético de *Drosophila*, cujos grandes cromossomos politênicos com bandas oferecem aos pesquisadores mapas citológicos com detalhes extraordinários.

LOCALIZAÇÃO DE GENES COM O AUXÍLIO DAS DELEÇÕES E DUPLICAÇÕES

Tomemos como exemplo de mapeamento citogenético as técnicas de localização do gene *white* ligado ao X de

Drosophila, uma cópia de tipo selvagem necessária para pigmentação dos olhos. Esse gene ocupa a posição 1,5 do mapa, perto de uma extremidade do cromossomo X. Mas está perto de qual das duas extremidades? E a que distância, em termos citológicos, está dessa extremidade? Para responder a essas perguntas, precisamos encontrar a posição do gene *white* no mapa citológico do cromossomo X politênico.

Um método é produzir moscas heterozigotas para uma mutação nula recessiva do gene *white* (*w*) e uma deleção (ou deficiência, geralmente simbolizada por *Df*) citologicamente definida em parte do cromossomo X (Figura 7.18). Esses heterozigotos *w/Df* servem como teste funcional para localizar o gene *white* em relação à deficiência. Se o gene *white* tiver sido deletado do cromossomo *Df*, então os heterozigotos *w/Df* não serão capazes de produzir o pigmento do olho porque não terão uma cópia ativa do gene *white* em nenhum dos cromossomos X. Portanto, os olhos dos heterozigotos *w/Df* serão brancos (o fenótipo mutante). Se, porém, o gene *white* não tiver sido deletado do cromossomo *Df*, os heterozigotos *w/Df* terão um gene *white* ativo em algum lugar desse cromossomo, e seus olhos serão vermelhos (o fenótipo selvagem). Portanto, o exame dos olhos dos heterozigotos *w/Df* torna possível determinar se uma deficiência específica deletou o gene *white*. Caso isso tenha acontecido, é preciso localizar o gene *white* dentro dos limites dessa deficiência.



■ FIGURA 7.18 Princípios de mapeamento de deleção para localizar um gene em cromossomo de *Drosophila*. O gene *white* no cromossomo X, definido pela mutação recessiva *w* que causa olhos brancos, é usado como exemplo.

Diferentes deficiências do cromossomo X possibilitaram que os pesquisadores localizassem o gene *white* perto da extremidade esquerda do cromossomo X (Figura 7.19). Cada deficiência foi combinada a uma mutação *white* recessiva, mas só uma das deficiências, *Df(1)w^{rJ1}*, produziu olhos brancos. Como essa deficiência “revela” a mutação *white*, sabemos que o gene *white* está obrigatoriamente localizado no segmento do cromossomo que ela deleta, isto é, entre as bandas 3A1 e 3C2 do cromossomo politênico. Em deficiências menores, o gene *white* foi localizado na banda 3C2 do cromossomo politênico, perto do limite direito de *Df(1)w^{rJ1}*.

Também podemos usar duplicações para identificar a localização citológica dos genes. O procedimento é semelhante ao que emprega deleções, exceto por pesquisarmos uma duplicação que mascara o fenótipo de uma mutação recessiva. A Figura 7.20 mostra um exemplo de uso de duplicações de pequenos segmentos do cromossomo X que foram translocados para outro cromossomo. Apenas uma dessas duplicações, *Dp2*, mascara – ou, como gostam de dizer os geneticistas, “cobre” – a mutação *white*; assim, ele tem obrigatoriamente uma cópia de tipo selvagem de *white*. Isso localiza o gene *white* entre as seções 2D e 3D no cromossomo X politênico, o que é compatível com os resultados dos testes de deleção já expostos.

As deleções e duplicações foram extraordinariamente úteis para localizar genes nos mapas citológicos de cromossomos de *Drosophila*. O princípio básico no mapeamento da deleção é que uma deleção que revela uma mutação recessiva não tem uma cópia de tipo selvagem do gene mutante. Esse fato localiza esse gene dentro dos limites da deleção. O princípio básico no mapeamento da duplicação é que uma duplicação que cobre uma mutação recessiva contém obrigatoriamente uma cópia de tipo selvagem do gene mutante. Esse fato localiza esse gene dentro dos limites da duplicação. Teste sua capacidade de localizar genes com base em deficiências e duplicações acompanhando o exercício do Boxe Resolva!: Mapeamento citológico de um gene de *Drosophila*.

Heterozigotos <i>w/Df</i>	Deficiência	Pontos de quebra	Fenótipo
	<i>Df(1)w^{rJ1}</i>	3A1; 3C2	Olhos brancos
	<i>Df(1)ct⁷⁸</i>	6F1-2; 7C1-2	Olhos vermelhos
	<i>Df(1)m²⁵⁹⁴</i>	10C1-2; 10E1-2	Olhos vermelhos
	<i>Df(1)r^{+75c}</i>	14B13; 15A9	Olhos vermelhos
	<i>Df(1)mal³</i>	19A1-2; 20A	Olhos vermelhos

A cor mutante dos olhos observada em *Df(1)w^{rJ1}* indica que o gene *white* está entre os pontos de quebra da deficiência nas bandas 3A1 e 3C2 no cromossomo X.

■ FIGURA 7.19 Localização do gene *white* no cromossomo X de *Drosophila* por mapeamento de deleção. Os pontos de quebra de deficiência são apresentados usando as coordenadas do mapa citológico de Bridges do cromossomo X politênico.

Combinações w/Dp	Duplicação	Pontos de quebra	Fenótipo
	Dp1	extremidade; 1E2-4	Olhos brancos
	Dp2	2D; 3D	Olhos vermelhos
	Dp3	6E2; 7C4-6	Olhos brancos
	Dp4	9F3; 10E3-4	Olhos brancos
	Dp5	14B13; 15A9	Olhos brancos

A cor dos olhos de tipo selvagem observada em Dp2 indica que o gene *white* está entre os pontos de quebra da duplicação nas regiões 2D e 3D no cromossomo X.

FIGURA 7.20 Localização do gene *white* no cromossomo X de *Drosophila* por mapeamento de duplicação. Cada duplicação é um segmento do cromossomo X que foi translocado para outro cromossomo. Para simplificar, porém, o outro cromossomo não é mostrado. Os pontos de quebra da duplicação são apresentados por meio das coordenadas do mapa citológico de Bridges do cromossomo X politênico.

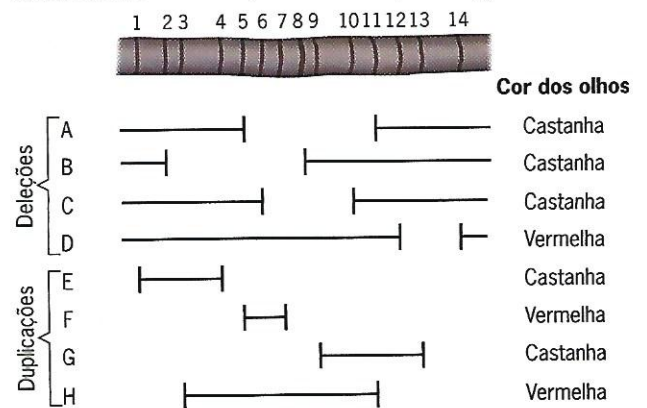
DISTÂNCIA GENÉTICA E DISTÂNCIA FÍSICA

Os procedimentos para medida da distância genética e construção de mapas de recombinação são baseados na incidência de *crossing over* entre cromossomos pareados. Intuitivamente, espera-se que os cromossomos longos tenham mais *crossing overs* que os curtos e que essa relação seja refletida no comprimento de seus mapas genéticos. Na maioria das vezes, nossa hipótese é verdadeira;

Resolva!

Mapeamento citológico de um gene de *Drosophila*

Uma mutação recessiva ligada ao X produz olhos castanhos em *Drosophila* hemizigota ou homozigota para ele; os olhos das moscas de tipo selvagem são vermelhos. Um geneticista produziu fêmeas que tinham essa mutação recessiva em um de seus cromossomos X; o outro cromossomo X tinha uma deleção citologicamente definida. O geneticista também produziu machos que tinham a mutação de olhos castanhos em seu cromossomo X; o cromossomo Y nesses machos tinha uma duplicação citologicamente definida de um pequeno segmento do cromossomo X. A extensão de cada deleção e duplicação é mostrada adiante em relação a um mapa de 14 bandas no cromossomo X politênico. Todas as fêmeas com mutação/deleção e todos os machos com mutação/duplicação foram classificados de acordo com a cor dos olhos. A partir dos resultados, localize o gene da cor do olho no menor intervalo possível no mapa citológico.



► Leia a resposta do problema no site <http://gen-io.grupogen.com.br>.

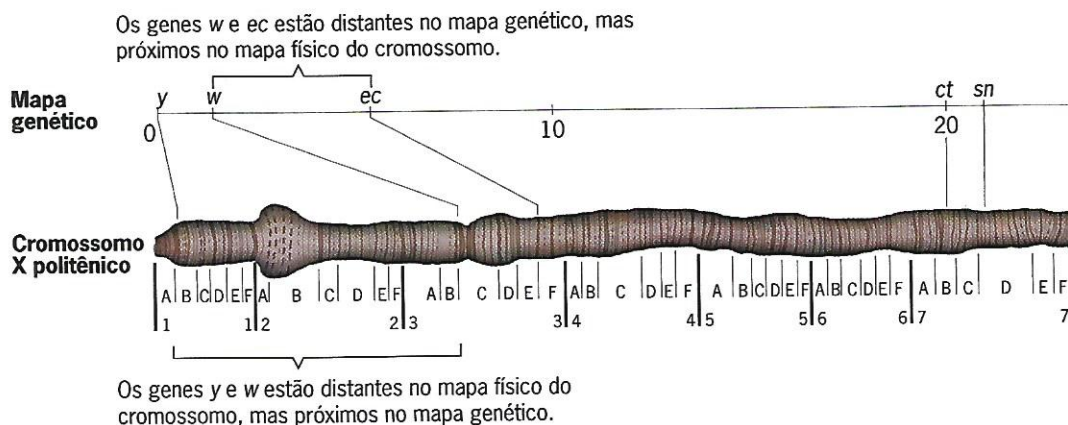


FIGURA 7.21 Extremidade esquerda do cromossomo X politênico de *Drosophila* e a porção correspondente do mapa genético mostrando os genes para corpo amarelo (*y*), olhos brancos (*w*), olhos equinos (*ec*), asas cortadas (*ct*) e cerdas chamuscadas (*sn*).

giões estão condensadas no mapa genético. Outras regiões, nas quais os *crossing overs* são mais frequentes, estão expandidas.

Embora não haja relação uniforme entre distância física e genética, os mapas genéticos e citológicos de um

cromossomo são colineares; isto é, locais específicos têm a mesma ordem. Portanto, o mapeamento de recombinação revela a verdadeira ordem dos genes ao longo de um cromossomo. No entanto, não mostra as distâncias físicas reais entre eles.

PONTOS ESSENCIAIS

- Em *Drosophila*, os genes podem ser localizados em mapas dos cromossomos politênicos por combinação de mutações recessivas com deleções e duplicações citologicamente definidas
- Uma deleção revela o fenótipo de uma mutação recessiva localizada entre seus limites, ao passo que uma duplicação oculta o fenótipo mutante
- Os mapas genéticos e citológicos são colineares; entretanto, as distâncias genéticas não são proporcionais às distâncias citológicas.

Análise de ligação em seres humanos

A análise do heredograma torna possível localizar genes nos cromossomos humanos.

Para detectar e analisar a ligação gênica em seres humanos, os geneticistas precisam colher dados de heredogramas. Muitas vezes esses dados são limitados ou incompletos, ou oferecem informações ambíguas. Portanto, a tarefa de elaborar mapas de ligação humana cria muitos desafios para os pesquisadores. Estudos clássicos de ligação em seres humanos concentraram-se em heredogramas nos quais era possível acompanhar simultaneamente a herança de dois genes ou mais. Hoje, graças aos métodos moleculares modernos, os pesquisadores podem analisar a herança de dezenas de marcadores diferentes no mesmo conjunto de heredogramas. Essa análise *multilocus* aumentou muito a capacidade de detectar a ligação e de elaborar mapas cromossômicos detalhados. As relações de ligação gênica mais fáceis de estudar em seres humanos são aquelas entre genes no cromossomo X. Esses genes seguem um padrão de herança prontamente identificado. Se dois genes tiverem esse padrão, eles estão ligados. É muito mais difícil identificar a ligação entre genes autossômicos. O genoma humano tem 22 autossomos diferentes, e um gene não ligado ao X poderia estar em qualquer um deles. Em que autossomo está o gene? Que outros genes estão ligados a ele? Quais são as posições desses genes no mapa? Essas são questões que desafiavam o pesquisador em genética humana.

Para saber como a ligação gênica é detectada em heredogramas humanos, examinemos parte do trabalho de J. H. Renwick e S. D. Lawler. Em 1955 esses pesquisadores relataram evidências de ligação entre o gene controlador do grupo sanguíneo do sistema ABO (Capítulo 4) e uma mutação dominante responsável por um distúrbio autossômico raro denominado síndrome unha-patela. As pessoas com essa síndrome têm anormalidade das unhas e das patelas. A Figura 7.22A mostra parte de um dos heredogramas estudados por Renwick e Lawler. Todos os indivíduos nesse heredograma foram caracterizados

quanto à presença ou ausência da mutação da síndrome unha-patela, designada *NPS1*; além disso, determinou-se o tipo sanguíneo ABO da maioria dos indivíduos.

A mulher da geração II representa uma nova ocorrência da mutação *NPS1*. Nenhum de seus pais nem de seus 11 irmãos tinha o fenótipo da síndrome unha-patela. Entre os cinco indivíduos que tinham a síndrome unha-patela nesse heredograma, todos eles, exceto um (III-6), tinham sangue tipo B. Essa observação sugere que a mutação *NPS1* está ligada geneticamente ao alelo *B* do *locus* do grupo sanguíneo ABO. Se presumirmos que essa inferência está correta, a mulher na geração II tem o genótipo *NPS1 B/+ O*; isto é, ela é um heterozigoto em repulsão. Sem dúvida, o genótipo de seu marido é *+ O/+ O*.

A Figura 7.22B ilustra os fenômenos genéticos desse heredograma e sugere uma estratégia para estimar, ainda que *grosso modo*, a distância entre os *loci NPS1* e *ABO*. O casamento indicado na Figura 7.22B é praticamente um cruzamento-teste. A mulher II-1 pode produzir quatro tipos diferentes de gametas, dois com cromossomos recombinantes e dois com cromossomos não recombinantes. A combinação desses gametas com o tipo simples de gameta (*+ O*) produzido pelo homem II-2 produz quatro genótipos diferentes. Como mostra o heredograma na Figura 7.22A, II-1 e II-2 tiveram todos os quatro tipos de filhos. No entanto, apenas 3 (III-3, III-6, III-12, indicados por asteriscos na Figura 7.22A) de seus 10 filhos eram recombinantes; os outros 7 eram não recombinantes. Assim, podemos estimar a frequência de recombinação entre os *loci NPS1* e *ABO* como $3/10 = 30\%$. No entanto, essa estimativa não usa todas as informações do heredograma. Para aprimorá-la, podemos incorporar as informações dos três netos dos casais, dos quais apenas um (IV-1) era recombinante. Ao todo, então, $3 + 1 = 4$ dos $10 + 3 = 13$ indivíduos da prole no heredograma eram recombinantes. Assim, concluímos que a frequência de recombinação entre os *loci NPS1* e *ABO* é de $4/13 = 31\%$. Em termos de um mapa de ligação, estimamos que a distância entre esses genes é de aproximadamente 31 cM. Renwick

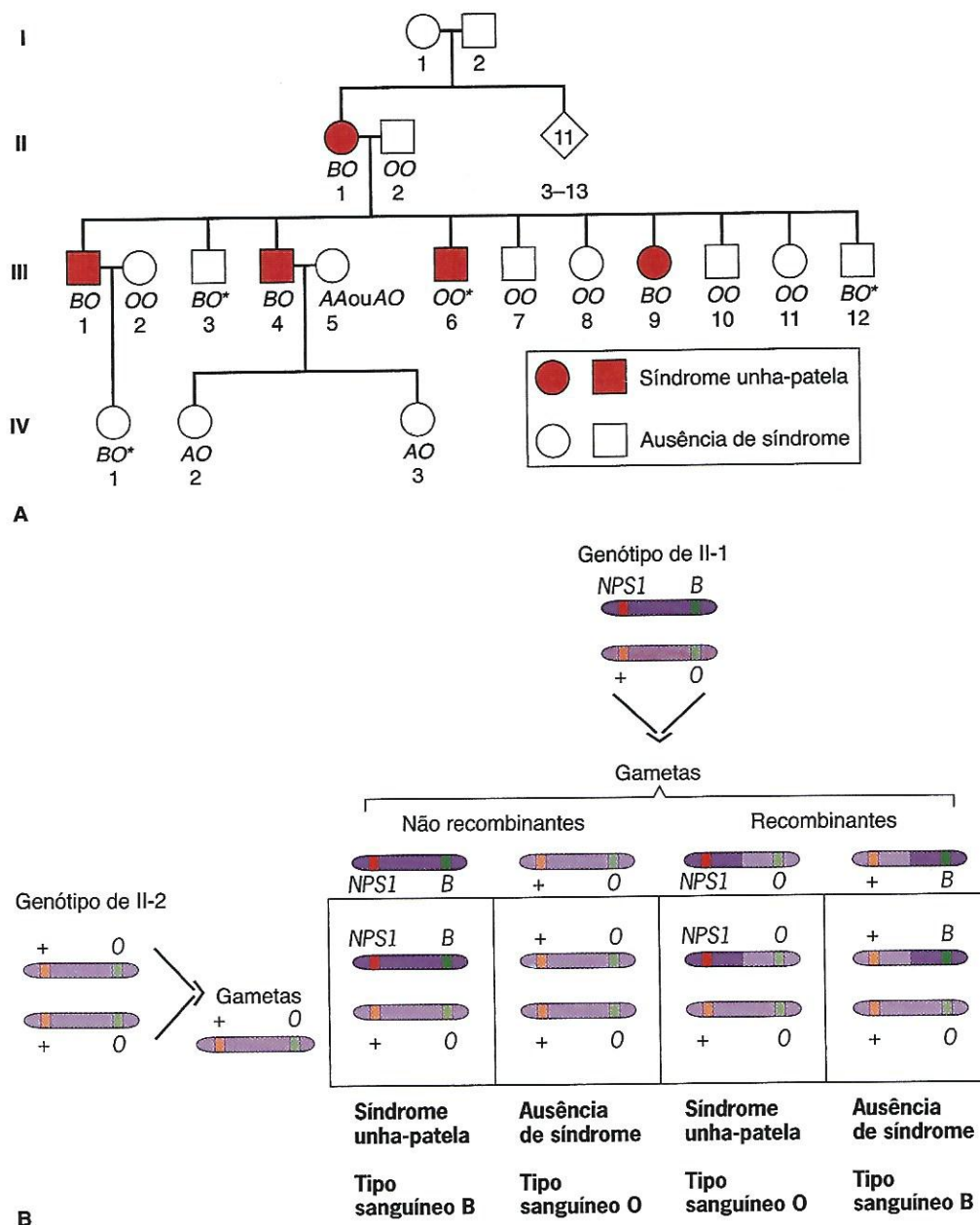


FIGURA 7.22 Análise da ligação gênica em heredograma humano. **A.** Parte de um heredograma mostrando a ligação entre os *loci* do sistema ABO e da síndrome unha-patela. Os indivíduos afetados pela síndrome unha-patela são indicados por símbolos vermelhos. Quando conhecido, o genótipo do *locus* ABO é apresentado abaixo de cada símbolo. Os asteriscos indicam os recombinantes. **B.** Quadrado de Punnett mostrando os genótipos produzidos pelo casal na geração II.

e Lawler analisaram a ligação entre os genes *NPS1* e *ABO* em outros heredogramas. Combinando todos os dados, eles estimaram que a frequência de recombinação é de aproximadamente 10%. Assim, a distância entre os genes *NPS1* e *ABO* é de aproximadamente 10 cM.

O estudo de Renwick e Lawler dos *loci* *NPS1* e *ABO* estabeleceu que esses dois genes estão ligados, mas não identificou o autossomo específico em que estão localizados. A primeira localização de um gene em um autossomo humano específico ocorreu em 1968, quando R. P. Donahue e colaboradores mostraram que o *locus* do grupo sanguíneo Duffy, designado *FY*, está no cromossomo 1. Essa demonstração baseou-se na descoberta de uma variante do cromossomo 1 que era mais longa que o

normal. A análise do heredograma mostrou que, em determinada família, esse cromossomo longo era segregado com alelos *FY* específicos. Assim, o *locus* *FY* foi localizado no cromossomo 1. A pesquisa subsequente localizou esse *locus* na região 1p31 desse cromossomo. Usando técnicas diferentes, os *loci* *NPS1* e *ABO* foram situados perto da extremidade do braço longo do cromossomo 9.

Até o início da década de 1980, o avanço no mapeamento gênico humano foi lentíssimo porque era difícil encontrar heredogramas com segregação de marcadores ligados, como duas doenças genéticas diferentes. Na década de 1980, porém, tornou-se possível identificar variantes genéticas no próprio DNA. Essas variantes resultam de diferenças na sequência de DNA em partes

dos cromossomos. Por exemplo, em um indivíduo, determinada sequência poderia ser GAATTC em um dos filamentos de DNA, e, em outro indivíduo, a sequência de DNA correspondente poderia ser GATTTC, uma diferença de apenas um nucleotídeo. Embora precisemos adiar para os capítulos posteriores a discussão das técnicas usadas para revelar essas diferenças moleculares, aqui podemos analisar como elas ajudaram a mapear os genes humanos, muitos deles implicados em doenças hereditárias graves. Se, além da análise fenotípica habitual, os membros de um heredograma são analisados quanto à presença ou ausência de marcadores moleculares no DNA, um pesquisador pode procurar a ligação entre cada marcador e o gene em estudo. Então, com técnicas estatísticas apropriadas, pode estimar as distâncias entre o gene e os marcadores ligados a ele.

Esse método permitiu que os geneticistas mapeassem um grande número de genes implicados em doenças humanas. Um dos exemplos mais expressivos é a pesquisa que localizou no cromossomo 4 o gene da doença de Huntington (*HD*), um distúrbio neurológico debilitante e, por fim, fatal. Nessa pesquisa, analisou-se a ligação entre o gene *HD* e um conjunto de marcadores moleculares em grandes heredogramas. À custa de trabalho

metodológico, o gene *HD* foi mapeado a 4 cM de um desses marcadores. Essa localização precisa estabeleceu os alicerces para o isolamento e a caracterização molecular do próprio gene *HD*.

Os marcadores moleculares também tornaram possível criar mapas de cromossomos humanos a partir de análises totalmente independentes. Caso se tenha demonstrado que o gene *A* está ligado ao marcador *x* em um conjunto de heredogramas e que o gene *B* está ligado ao marcador *x* em outro conjunto de heredogramas, então é óbvio que existe uma ligação entre o gene *A* e o gene *B*. Assim, a análise desses marcadores facilita aos geneticistas humanos determinar as relações de ligação entre genes que não são segregados nos mesmos heredogramas.

A análise de dados de recombinação dos heredogramas possibilita que os geneticistas elaborem mapas de ligação gênica dos cromossomos. Contudo, exceto no caso de ligação ao X, essa análise não nos diz que cromossomo está sendo mapeado, ou onde está determinado gene na imagem física desse cromossomo. Esses desafios foram enfrentados pelo desenvolvimento de técnicas citológicas como bandeamento cromossômico e pintura cromossômica (Capítulo 6).

PONTOS ESSENCIAIS

- A ligação entre genes humanos pode ser detectada por análise dos heredogramas
- A análise do heredograma também possibilita estimar as frequências de recombinação para mapear genes nos cromossomos humanos.

Recombinação e evolução

A recombinação – ou sua ausência – tem um papel essencial na evolução.

A recombinação é uma característica essencial da reprodução sexuada. Durante a meiose, por ocasião da aproximação dos cromossomos e do *crossing over*, há uma oportunidade de criar novas combinações de alelos. Algumas delas podem beneficiar o organismo por aumento da sobrevivência ou da capacidade reprodutiva. Com o tempo, o esperado é que essas combinações benéficas se disseminem na população e se tornem características usuais da constituição genética da espécie. Portanto, a recombinação meiótica é uma maneira de embaralhar a variação genética para potencializar as mudanças evolutivas.

SIGNIFICADO EVOLUTIVO DA RECOMBINAÇÃO

É possível avaliar a vantagem evolutiva da recombinação mediante comparação de duas espécies, uma capaz de se reproduzir de maneira sexuada e outra, não. Suponhamos que tenha surgido uma mutação benéfica em cada

espécie. Ao longo do tempo, o esperado seria a disseminação dessas mutações. Suponhamos também que, enquanto elas estão se disseminando, ocorra outra mutação benéfica em um indivíduo não mutante de cada espécie. No organismo assexuado, não há possibilidade de que essa segunda mutação seja recombinada com a primeira, mas, no organismo sexuada, as duas mutações podem ser recombinadas com a produção de uma linhagem melhor que qualquer um dos mutantes isolados. Essa linhagem recombinante será capaz de se disseminar por toda a população da espécie. Em termos evolutivos, a recombinação pode permitir a reunião de alelos favoráveis de diferentes genes no mesmo organismo.

SUPRESSÃO DA RECOMBINAÇÃO POR INVERSÕES

O efeito de embaralhamento de genes na recombinação pode ser impedido por rearranjos do cromossomo. O *crossing over* geralmente é inibido perto dos pontos de quebra de um rearranjo em condição heterozigota, provavelmente porque o rearranjo desorganiza o pare-

amento dos cromossomos. Portanto, muitos rearranjos estão associados à redução na frequência de recombinação. Esse efeito é mais acentuado em heterozigotos por inversão porque a inibição do *crossing over* que ocorre perto dos pontos de quebra da inversão é complicada pela perda seletiva de cromossomos que tiveram *crossing over* na região invertida.

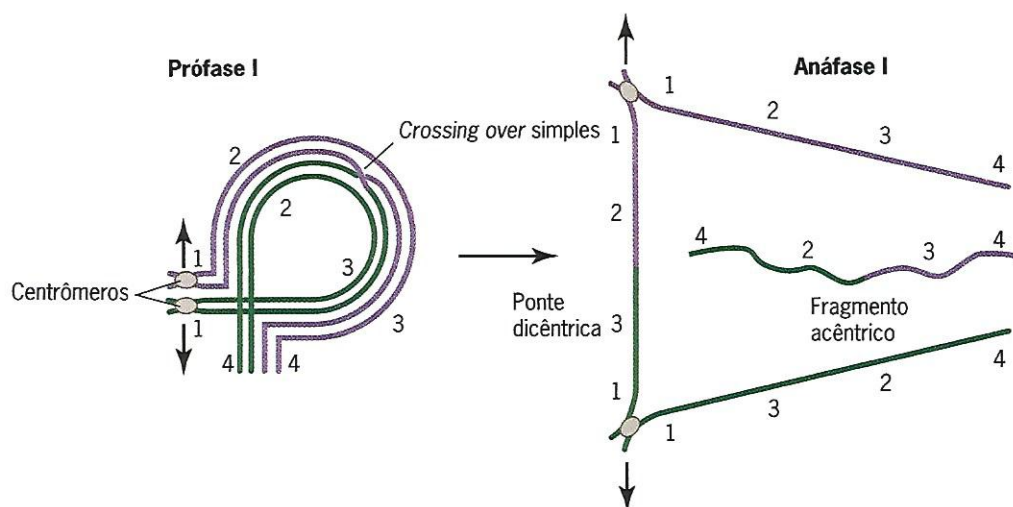
Para analisar esse efeito de supressão da recombinação, consideremos uma inversão no braço longo de um cromossomo (Figura 7.23). O *crossing over* entre cromátides invertidas e não invertidas na tétrade produz duas cromátides recombinantes, mas provavelmente as duas cromátides serão perdidas durante a meiose ou depois dela. Uma das cromátides não tem centrômero – é um fragmento *acêntrico* – e, portanto, é incapaz de se deslocar até a devida posição durante a anáfase da primeira divisão meiótica. A outra cromátide tem dois centrômeros e, portanto, será puxada em direções opostas, formando uma *ponte de cromátide dicêntrica*. Por fim, essa ponte se quebra e divide a cromátide em pedaços. Ainda que as cromátides acêntrica e dicêntrica produzidas por *crossing over* na inversão sobrevivam à meiose, é improvável que os zigotos sejam viáveis. As duas cromátides são aneuploides – com duplicação de alguns genes e deficiência de outros – e a aneuploidia geralmente é letal. Portanto, essas cromátides serão eliminadas por seleção natural na próxima geração. O efeito final dessa perda de cromátides é a supressão da recombinação entre cromossomos invertidos e não invertidos em heterozigotos.

Os geneticistas exploraram as propriedades de supressão da recombinação pelas inversões para manter alelos de diferentes genes juntos no mesmo cromossomo. Suponhamos, por exemplo, que um cromossomo cuja estrutura é normal tenha os alelos recessivos *a*, *b*, *c*, *d* e *e*. Se houver pareamento desse cromossomo com outro de estrutura normal que tenha os alelos selvagens correspondentes *a*⁺, *b*⁺, *c*⁺, *d*⁺ e *e*⁺, os alelos recessivos e

selvagens serão misturados por recombinação. Para evitar essa mistura, o cromossomo com os alelos recessivos pode ser pareado com o cromossomo de tipo selvagem que tem uma inversão. A menos que haja *crossing overs* duplos na região invertida, essa heterozigosidade estrutural suprimirá a recombinação. O cromossomo mutante multiplicado então pode ser transmitido para a prole na forma de uma unidade genética intacta.

Muitas vezes, essa técnica de supressão da recombinação foi usada em experimentos com *Drosophila*, na qual o cromossomo invertido geralmente tem uma mutação dominante que permite seu acompanhamento durante uma série de cruzamentos sem exames citológicos. Esse cromossomo marcado por inversão é denominado **balanceador**, pois permite manter um cromossomo de interesse em condição heterozigota sem quebra e recombinação.

A supressão da recombinação por inversão parece ter tido um papel importante na evolução dos cromossomos sexuais em mamíferos. Os dados provêm das análises de Bruce Lahn e David Page, que estudaram 19 genes presentes nos cromossomos X e Y humano. Esses genes em comum ocupam posições diferentes nos cromossomos X e Y, indicando que foram rearranjados por inversões durante a evolução. Além disso, as sequências de DNA das cópias ligadas ao X e ao Y desses genes em comum divergiram em diferentes graus. Pela análise da variação no grau de divergência, Lahn e Page distinguiram quatro “níveis evolutivos” nos cromossomos sexuais humanos – regiões nas quais a recombinação foi suprimida por diferentes períodos durante a evolução. Lahn e Page presumem que os cromossomos X e Y se originaram de um par de autossomos algum tempo depois que a linha evolutiva dos mamíferos divergiu da linha de antigos répteis que levou ao surgimento dos dinossauros, crocodilos e aves. Entre 240 e 320 milhões de anos atrás, uma inversão no que deveria se tornar o cromossomo Y causou a supressão regional da recombinação entre



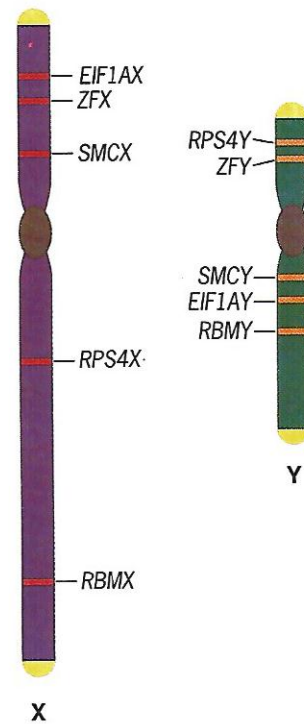
■ FIGURA 7.23 Supressão da recombinação em heterozigoto para inversão. Os cromossomos dicêntricos (1 2 3 1) e acêntricos (4 3 2 4) formados a partir das cromátides que tiveram *crossing over* são aneuploides e causam inviabilidade da geração subsequente. Consequentemente, os produtos do *crossing over* entre os cromossomos invertidos e não invertidos não são restabelecidos.

X e Y. Na linhagem que originou os seres humanos, houve no mínimo três outras inversões, duas delas em algum momento entre 80 e 130 milhões de anos atrás e uma delas entre 30 e 50 milhões de anos atrás. O efeito final dessas inversões foi suprimir a recombinação entre a maioria das regiões nos cromossomos X e Y. Por força da seleção natural, genes ativos foram preservados no cromossomo X, mas no cromossomo Y houve degeneração da maioria dos genes pelo acúmulo de mutações aleatórias. Assim, hoje o cromossomo Y tem muito menos genes ativos que o cromossomo X, e os genes remanescentes estão arrançados em outra ordem (Figura 7.24).

CONTROLE GENÉTICO DA RECOMBINAÇÃO

Não surpreende que um processo tão importante quanto a recombinação esteja sob controle genético. Estudos com vários organismos, entre eles leveduras e *Drosophila*, mostraram que produtos de muitos genes participam da recombinação. Alguns desses produtos gênicos participam do pareamento dos cromossomos, outros catalisam o processo de permuta, e ainda outros ajudam a reunir os segmentos das cromátides quebradas. Abordaremos algumas dessas atividades com mais detalhes no Capítulo 13.

Um fenômeno curioso, ainda inexplicado, é a não ocorrência de *crossing over* em machos de *Drosophila*. Nesse aspecto, a *Drosophila* é diferente da maioria das espécies,



■ FIGURA 7.24 Ordem de genes compartilhados fora das regiões pseudoautosômicas nos cromossomos X e Y humanos.

inclusive da nossa, nas quais o *crossing over* ocorre em ambos os sexos. Além disso, sabemos que o grau de recombinação varia entre as espécies. Talvez os próprios eventos que causam recombinação estejam sujeitos a mudanças evolutivas.

PONTOS ESSENCIAIS

- A recombinação pode reunir mutações favoráveis
- Os rearranjos cromossômicos, sobretudo as inversões, podem suprimir a recombinação
- A recombinação está sujeita a controle genético.

Exercícios

Aplice a análise genética básica

- Uma linhagem endogâmica de boca-de-leão de flores roxas e folhas foscas foi cruzada com outra linhagem endogâmica de flores brancas e folhas brilhantes. As plantas da F_1 , todas as quais tinham flores roxas e folhas foscas, foram retrocruzadas com a linhagem de flores brancas e folhas brilhantes, e obtiveram-se as seguintes plantas da F_2 : 50 com flores roxas e folhas foscas; 46 com flores brancas e folhas brilhantes; 12 com flores roxas e folhas brilhantes e 10 com flores brancas e folhas foscas. (a) Quais das quatro classes da F_2 são recombinantes? (b) Qual é a evidência da ligação entre os genes para cor das flores e textura das folhas?

(c) Faça o diagrama dos cruzamentos desse experimento. (d) Qual é a frequência de recombinação entre os genes que determinam a cor das flores e a textura das folhas? (e) Qual é a distância de mapa genético entre esses genes?

Resposta: (a) As duas últimas classes – flores roxas com folhas brilhantes e flores brancas com folhas foscas – na F_2 são recombinantes. Nenhuma dessas combinações de fenótipos estava presente nas linhagens usadas no cruzamento inicial. (b) Os recombinantes são 18,6% das plantas da F_2 – muito menos que os 50% que seriam esperados se os genes da cor das flores e da textura das folhas

não estivessem ligados. Portanto, esses genes estão ligados no mesmo cromossomo no genoma da boca-de-leão. (c) Para representar os cruzamentos, primeiro é preciso atribuir símbolos aos alelos dos genes da cor das flores e da textura das folhas: W = roxa, w = branca; S = fosca, s = brilhante; a letra maiúscula indica que o alelo é dominante. O primeiro cruzamento é $WS/WS \times w s/w s$, produzindo plantas da F_1 com o genótipo $WS/w s$. O retrocruzamento é $WS/w s \times w s/w s$, produzindo quatro classes de prole: (1) $WS/w s$, (2) $w s/w s$, (3) $W s/w s$ e (4) $w S/w s$. As classes 1 e 2 são tipos parentais, e as classes 3 e 4 são recombinantes. (d) A frequência de recombinação é 18,6%. (e) A distância de mapa genético é estimada pela frequência de recombinação como 18,6 centiMorgans.

2. Qual é a indicação citológica de que houve *crossing over*? Quando e onde você procuraria por ele?

Resposta: É provável que o *crossing over* ocorra do início ao meio da prófase da meiose I. Não é fácil, porém, analisar os cromossomos nesses estágios, e é difícil, se não impossível, identificar as permutas por métodos citológicos. A melhor evidência citológica de que houve *crossing over* é obtida das células perto do fim da prófase da meiose I. Nesse estágio, os homólogos pareados repelem-se levemente, e as permutas entre eles são observadas como quiasmas.

3. Um geneticista estimou o número de permutas ocorridas durante a meiose em cada uma das 100 cromátides distinguidas em gametas. Os dados são:

Número de permutas	Frequência
0	18
1	20
2	40
3	16
4	6

Qual é o comprimento genético, em centiMorgans, do cromossomo analisado nesse estudo?

Resposta: O comprimento genético de um cromossomo é o número médio de permutas em uma cromátide no fim da meiose. Para os dados apresentados, a média é $0 \times (18/100) + 1 \times (20/100) + 2 \times (40/100) + 3 \times (16/100) + 4 \times (6/100) = 1,72$ Morgans ou 172 centiMorgans.

4. Fêmeas de *Drosophila* heterozigotas para três marcadores recessivos ligados ao X, y (corpo amarelo), ct (asas cortadas) e m (asas em miniatura), e seus alelos selvagens foram cruzadas com machos $y ct m$. Obteve-se a seguinte prole:

Classe fenotípica	Número
1. corpo amarelo, asas cortadas e em miniatura	30
2. tipo selvagem	33
3. corpo amarelo	10
4. asas cortadas e em miniatura	12
5. asas em miniatura	8
6. corpo amarelo, asas cortadas	5
7. corpo amarelo, asas em miniatura	1
8. asas cortadas	1
Total:	100

(a) Que classes são tipos parentais? (b) Que classes representam *crossing overs* duplos? (c) Que gene está no meio dos outros dois? (d) Qual era o genótipo das fêmeas heterozigotas usadas no cruzamento? (Indique a fase de ligação correta, bem como a ordem correta dos marcadores ao longo do cromossomo.)

Resposta: (a) As classes parentais são as mais numerosas; portanto, nesses dados, as classes 1 e 2 são tipos parentais. (b) As classes com *crossing over* duplo são as menos numerosas; portanto, nesses dados, as classes 7 e 8 são as classes com *crossing over* duplo. (c) As classes parentais mostram que os três alelos mutantes entraram nas fêmeas heterozigotas no mesmo cromossomo X; o outro cromossomo X nessas fêmeas tinha, obrigatoriamente, os três alelos selvagens. As classes de *crossing over* duplo informam qual dos três genes está no meio porque o marcador intermediário é separado dos marcadores flancadores pelo processo de permuta dupla. Nesses dados, o alelo ct está separado de y e m nas classes de *crossing over* duplo; portanto, o gene ct tem de estar entre os genes y e m . (d) O genótipo das fêmeas heterozigotas usadas no cruzamento era, obrigatoriamente, $y ct m/+ + +$.

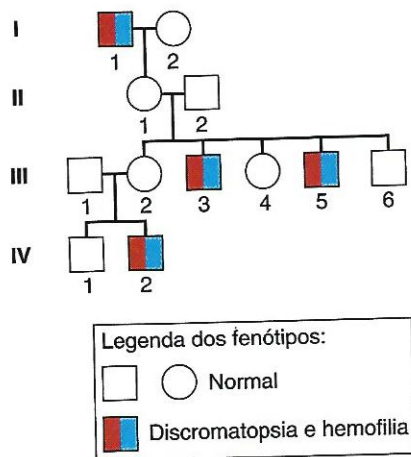
5. Um geneticista de *Drosophila* fez experimentos para localizar o gene das cerdas chamuscadas (*singed* [sn]) no mapa citológico do cromossomo X. Machos hemizigotos para uma mutação sn recessiva foram cruzados com fêmeas que tinham várias deficiências (indicadas por Df) no cromossomo X balanceadas em relação a um cromossomo X com múltiplas inversões marcado pela mutação semidominante para olhos em barra (Bar [B]). Assim, o esquema de cruzamento era machos $sn/Y \times$ fêmeas Df/B . Os resultados dos cruzamentos com quatro deficiências diferentes são:

Deficiência	Pontos de quebra	Fenótipo das filhas sem olhos em barra
1	2F; 3C	tipo selvagem
2	4D; 5C	tipo selvagem
3	6F; 7E	cerdas chamuscadas
4	7C; 8C	cerdas chamuscadas

O mapa citológico do cromossomo X é dividido em 20 seções numeradas, cada uma delas subdividida em subseções A a F. Onde está o gene *singed* (cerdas chamuscadas) neste mapa citológico?

Resposta: As filhas sem olhos em barra examinadas à procura do fenótipo cerdas chamuscadas tinham genótipo *Df/sn*. A mutação *singed* foi “revelada” por duas deficiências, 3 e 4; portanto, tem de estar na região deletada no cromossomo X comum a ambas, isto é, na região 7C–7E.

6. O heredograma a seguir mostra quatro gerações de uma família descrita em 1928 por M. Madlener. O bisavô, I-1, tem discromatopsia e hemofilia. Considerando que *c* representa o alelo para discromatopsia e *h*, o alelo para hemofilia, quais são os genótipos dos cinco netos desse homem? Algum indivíduo no heredograma apresenta evidências de recombinação entre os genes para discromatopsia e hemofilia?



Resposta: Os genes para discromatopsia e hemofilia são ligados ao X. Como I-1 tem discromatopsia e hemofilia, seu genótipo tem de ser *ch*. A filha, II-1, tem fenótipo normal e, portanto, tem, obrigatoriamente, os alelos não mutantes, *C* e *H*, desses dois genes ligados ao X. Além disso, como II-1 herdou *c* e *h* do pai, os dois alelos não mutantes existentes no seu genótipo têm de estar no cromossomo X herdado da mãe. Portanto, o genótipo de II-1 é *CH/cb*, isto é, ela é um heterozigoto em acoplamento para os dois loci. III-2, a primeira neta de I-1, também é um heterozigoto em acoplamento. Inferimos que ela tem esse genótipo porque seu filho tem discromatopsia e hemofilia (*ch*) e seu pai tem fenótipo normal (*CH*). É evidente que III-2 herdou o cromossomo *ch* da mãe. Entre os netos de sexo masculino de I-1, dois (III-3 e III-5) têm tanto hemofilia quanto discromatopsia; assim, esses netos têm genótipo *ch*. O outro neto de sexo masculino (III-6) não tem discromatopsia nem hemofilia; portanto, seu genótipo é *CH*. O genótipo da neta remanescente (III-4) é incerto. Ela herdou um cromossomo *CH* do pai, mas o cromossomo herdado da mãe pode ser *CH*, *ch*, *Ch*, ou *cH*. Não é possível determinar pelo heredograma qual desses cromossomos ela recebeu. O máximo que podemos dizer sobre o genótipo de III-4 é que há um cromossomo com os alelos *C* e *H*.

Nenhum dos quatro netos aos quais podemos atribuir genótipos apresenta evidências de recombinação entre os genes para discromatopsia e hemofilia, assim como nenhum dos dois bisnetos da geração IV. Um desses bisnetos tem genótipo *CH* e o outro, *ch*. Assim, no heredograma como um todo não existem evidências de recombinação entre os genes *C* e *H*.

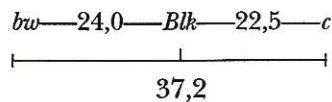
Autoavaliação

Integre diferentes conceitos e técnicas

1. R. K. Sakai, K. Akhtar e C. J. Dubash (1985, *J. Hered.* 76:140-141) apresentaram dados de um grupo de cruzamentos-teste com o mosquito *Anopheles culicifacies*, um vetor da malária no sul da Ásia. Os dados apontavam três mutações: *bw* (olhos castanhos [*brown*]), *c* (olhos incolores [*colorless*]) e *Blk* (corpo preto [*black*]). Em cada cruzamento, heterozigotos para repulsão foram cruzados com mosquitos homozigotos para os alelos recessivos dos genes, e a prole foi classificada em genótipo parental ou recombinante. Há alguma ligação entre esses três genes estudados nos cruzamentos? Em caso afirmativo, elabore um mapa das relações de ligação.

Cruzamento	Heterozigoto para repulsão	Prole		
		Parental	Recombinante	Porcentagem de recombinação
1	<i>bw +/+ c</i>	850	503	37,2
2	<i>bw +/+ Blk</i>	750	237	24,0
3	<i>c +/+ Blk</i>	629	183	22,5

Resposta: Em cada cruzamento, a frequência de recombinação é inferior a 50%, portanto os três loci estão ligados. Para representá-los em um mapa de ligação, estimamos as distâncias entre cada par de genes a partir das frequências de recombinação observadas:



Observe que a frequência de recombinação entre *bw* e *c* (37,2%, do cruzamento 1) é bem menor que a distância real entre esses genes (46,5). Isso mostra que, no caso de genes muito separados, a frequência de recombinação subestima a verdadeira distância de mapa.

2. Cerdas chamuscadas (*sn*), asas sem nervuras transversais (*cv*), e cor vermelhão dos olhos (*v*) são causadas por alelos mutantes recessivos de três genes ligados ao X em *Drosophila melanogaster*. No cruzamento-teste de uma fêmea heterozigota para todos esses três genes com um macho de cerdas chamuscadas, asas sem nervuras transversais e olhos vermelhão, obteve-se a seguinte prole:

Classe	Fenótipo	Número
1	cerdas chamuscadas, asas sem nervuras transversais, olhos vermelhão	3
2	asas sem nervuras transversais, olhos vermelhão	392
3	olhos vermelhão	34
4	asas sem nervuras transversais	61
5	cerdas chamuscadas, asas sem nervuras transversais	32
6	cerdas chamuscadas, olhos vermelhão	65
7	cerdas chamuscadas	410
8	tipo selvagem	3
		Total: 1.000

Qual é a ordem correta desses três genes no cromossomo X? Quais são as distâncias de mapa genético entre *sn* e *cv*, *sn* e *v*, e *cv* e *v*? Qual é o coeficiente de coincidência?

Resposta: Antes de tentar analisar esses dados, é preciso determinar o genótipo da fêmea heterozigota que produziu as oito classes de prole. Para fazer isso, identificamos as duas classes parentais (2 e 7), que são as mais numerosas. Essas classes informam que a fêmea heterozigota tinha as mutações *cv* e *v* em um dos cromossomos X e a mutação *sn* no outro. Portanto, o genótipo era (*cv* + *v*) / (+ *sn* +), escritos entre parênteses para indicar incerteza acerca da ordem dos genes.

Para determinar a ordem dos genes, é preciso identificar as classes com *crossing over* duplo entre os seis tipos de prole recombinante. Essas são as classes 1 e 8, as menos numerosas. Elas mostram que o gene *singed* está entre os genes *crossveinless* e *vermilion*. Podemos verificar isso investigando o efeito de um *crossing over* duplo em uma fêmea que tem o genótipo.

$$\frac{cv + v}{+ sn +}$$

Duas permutas nesse genótipo produzem gametas *cv sn v* ou + + +, que correspondem às classes 1 e 8, os *crossing overs* duplos observados. Assim, a ordem proposta dos genes – *cv sn v* – está correta.

Estabelecida a ordem dos genes, agora é possível identificar quais classes recombinantes representam *crossing overs* entre *cv* e *sn*, e quais representam *crossing overs* entre *sn* e *v*.

Crossing overs entre *cv* e *sn*:

Classe:	3	5	1	8	
Número:	34	+ 32	+ 3	+ 3	= 72

Crossing overs entre *sn* e *v*:

Classe:	4	6	1	8	
Número:	61	+ 65	+ 3	+ 3	= 132

Determinamos as distâncias entre esses pares de genes calculando o número médio de *crossing overs*. Entre *cv* e *sn*, a distância é 72/1.000 = 7,2 cM, e entre *sn* e *v* é 132/1.000 = 13,2 cM. Podemos estimar a distância entre *cv* e *v* como a soma desses valores: 7,2 + 13,2 = 20,4 cM. Portanto, o mapa de ligação desses três genes é:

$$cv \text{---} 7,2 \text{---} sn \text{---} 13,2 \text{---} v$$

Para calcular o coeficiente de coincidência, usamos as frequências observada e esperada de *crossing overs* duplos:

$$c = \frac{\text{frequência observada de } \textit{crossing overs} \text{ duplos}}{\text{frequência esperada de } \textit{crossing overs} \text{ duplos}} = \frac{0,006}{0,072 \times 0,132} = 0,63$$

que indicam apenas interferência moderada.

3. Uma geneticista de *Drosophila* está pesquisando uma mutação letal recessiva, *l(1)r13*, localizada no cromossomo X. Essa mutação é mantida em estoque com um cromossomo X balanceador marcado com uma mutação semidominante para olhos em barra (*B*). Em condição homozigota e hemizigota, a mutação *B* reduz os olhos a barras estreitas. Na condição heterozigota, causa olhos reniformes. As moscas homozigotas ou hemizigotas para o alelo selvagem de *B* têm olhos grandes e esféricos. Para manter a mutação *l(1)r13* em estoque, a cada geração a geneticista cruza machos *B* com fêmeas *l(1)r13/B* e seleciona as filhas de olhos reniformes para cruzamento com seus irmãos de olhos em barra. O objetivo da geneticista é determinar a localização citológica de *l(1)r13*. Para isso, ela cruza fêmeas *l(1)r13/B* com vários machos que têm duplicações de segmentos curtos do cromossomo X em seus genomas. Todas as duplicações estão vinculadas ao cromossomo Y. Assim, o genótipo dos machos usados nesses cruzamentos pode ser representado como *X/Y-Dp*. A geneticista examina a prole de cada cru-

zamento à procura de machos sem olhos em barra. A partir dos resultados apresentados na tabela a seguir, determine a localização citológica de $l(1)r13$.

Nome Dp	Segmento Dp*	Presença de machos sem olhos em barra
1	2D–3D	Sim
2	3A–3E	Sim
3	3D–4A	Não
4	4A–4D	Não
5	4B–4E	Não

*O braço longo do cromossomo X é dividido em 20 seções numeradas, começando com a seção 1 na extremidade e terminando com a seção 20 perto do centrômero. Cada seção é subdividida em seis subseções, ordenadas alfabeticamente de A a F. A subseção A está no lado de cada seção numerada mais perto da extremidade.

Resposta: O cruzamento para manter a mutação letal em estoque é machos $B/Y \times$ fêmeas $l(1)r13/B \rightarrow$ machos B/Y (olhos em barra), machos $l(1)r13/Y$ (morrem), fêmeas $l(1)r13/B$ (olhos reniformes) e fêmeas B/B (olhos em barra). A cada geração, os machos B/Y e as fêmeas $l(1)r13/B$ são selecionados para cruzamentos a fim de perpetuar a mutação letal. Um cruzamento para determinar a localização citológica da mutação letal pode ser representado como fêmeas $l(1)r13/B \times$ machos $X/YDp \rightarrow$ machos $l(1)r13/YDp$ (se viáveis, sem olhos em barra), machos B/YDp (olhos em barra), fêmeas $l(1)r13/X$ (sem olhos em barra) e fêmeas B/X (olhos reniformes). A primeira classe de moscas – machos sem olhos em barra – informa se uma duplicação específica “cobre” ou não a mutação letal. Se isso acontecer, esses machos aparecerão na prole da criação. Caso contrário, eles não aparecerão. A partir desses dados, verificamos que duas duplicações, $Dp\ 1$ e $Dp\ 2$,

cobrem a mutação letal. Assim, a mutação tem de estar dentro dos limites dessas duplicações – isto é, entre 2D e 3E. Podemos refinar a localização da mutação letal observando que há superposição das duas duplicações desde a subseção 3A até a subseção 3D. Portanto, a mutação está obrigatoriamente na região 3A–3D do cromossomo X.

4. Uma mulher tem duas características dominantes, cada uma delas causada por uma mutação em um gene diferente: catarata (uma anormalidade oftálmica), por herança paterna, e polidactilia (um dedo extranumerário), por herança materna. O marido não tem nenhuma dessas características. Se os genes dessas duas características estiverem 15 cM distantes no mesmo cromossomo, qual é a chance de que o primeiro filho desse casal tenha catarata e polidactilia?

Resposta: Para calcular a chance de que a criança tenha as duas características, primeiro é preciso determinar a fase de ligação dos alelos mutantes no genótipo da mulher. Como ela herdou a mutação para catarata do pai e a mutação para polidactilia da mãe, os alelos mutantes têm de estar em cromossomos opostos, isto é, na fase de ligação de repulsão:

$$\frac{C}{+} \frac{+}{P}$$

Para que uma criança herde os dois alelos mutantes, a mulher terá que produzir um ovócito com um cromossomo recombinante, $C\ P$. É possível estimar a probabilidade desse evento a partir da distância entre os dois genes, 15 cM, que, tendo em vista a interferência, deve ser equivalente à recombinação de 15%. No entanto, apenas metade dos recombinantes será $C\ P$. Assim, a chance de que a criança herde os dois alelos mutantes é de $(15/2)\% = 7,5\%$.

Avaliação adicional

Entenda melhor e desenvolva a capacidade analítica

- 7.1 Mendel desconhecia a existência de cromossomos. Caso conhecesse, que modificação poderia ter feito no princípio da distribuição independente?
- 7.2 O cruzamento entre indivíduos com os genótipos $Cc\ Dd\ Ee \times cc\ dd\ ee$ produziu uma prole de 1.000 organismos. A classe $C- D- ee$ tinha 351 indivíduos. Os genes c , d e e estão no mesmo cromossomo ou em diferentes cromossomos? Explique.
- 7.3 Se a está ligado a b , e b a c , e c a d , um experimento de recombinação detectaria ligação gênica entre a e d ? Explique.
- 7.4 Os camundongos têm 19 autossomos no genoma, todos com quase o mesmo tamanho. Qual é a chance de que dois genes autossômicos escolhidos aleatoriamente estejam no mesmo cromossomo?
- 7.5 Genes em cromossomos diferentes recombinam-se com uma frequência de 50%. É possível que dois genes no mesmo cromossomo se recombinem com essa frequência?
- 7.6 Se dois loci estiverem distantes 10 cM, que proporção das células na prófase da primeira divisão meiótica conterá um *crossing over* simples na região entre eles?
- 7.7 Os genes a e b estão distantes 20 cM. Um indivíduo $a^+ b^+ / a^+ b^+$ foi cruzado com outro $a b / a b$.
- (a) Represente o cruzamento em um diagrama e mostre os gametas produzidos por cada genitor e o genótipo da F_1 .

- (b) Que gametas os indivíduos da F_1 podem produzir? Em que proporções?
- (c) Qual seria a prole esperada do cruzamento da F_1 com indivíduos $a b/a b$? Em que proporções?
- (d) Esse é um exemplo da fase de ligação de acoplamento ou repulsão?
- (e) Qual seria a prole esperada do intercruzamento da F_1 ? Em que proporções?
- 7.8** Responda às perguntas (a) a (e) do problema anterior supondo que o cruzamento original tenha sido $a^+ b/a^+ b \times a b^+/a b^+$.
- 7.9** Se a frequência de recombinação nos dois problemas anteriores fosse de 40% em vez de 20%, que alteração ocorreria na proporção de gametas e da prole do cruzamento-teste?
- 7.10** Uma variedade homozigota de milho de folhas vermelhas e sementes normais foi cruzada com outra variedade homozigota de folhas verdes e sementes no pendão. Em seguida, os híbridos foram retrocruzados com a variedade de folhas verdes e sementes no pendão, e obteve-se a seguinte prole: 124 de folhas vermelhas e sementes normais; 126 de folhas vermelhas e sementes no pendão; 125 de folhas verdes e sementes normais; 123 de folhas verdes e sementes no pendão. Há ligação entre os genes que determinam a cor da planta e o tipo de semente? Explique.
- 7.11** Uma fêmea de tipo selvagem da mosca-das-frutas heterozigota para genes que controlam a cor do corpo e o comprimento das asas foi cruzada com um macho mutante homozigoto de corpo preto (alelo b) e asas vestigiais (alelo vg). O cruzamento produziu a seguinte prole: 126 de corpo cinza e asas normais; 24 de corpo cinza e asas vestigiais; 26 de corpo preto e asas normais; 124 de corpo preto e asas vestigiais. Esses dados indicam ligação entre os genes para cor do corpo e comprimento da asa? Qual é a frequência de recombinação? Represente o cruzamento, mostrando o arranjo dos marcadores genéticos nos cromossomos.
- 7.12** Outra fêmea de mosca-das-frutas com fenótipo selvagem heterozigota para os dois genes mencionados no problema anterior foi cruzada com um macho homozigoto de corpo preto e asas vestigiais. O cruzamento produziu a seguinte prole: 23 de corpo cinza e asas normais; 127 de corpo cinza e asas vestigiais; 124 de corpo preto e asas normais; 26 de corpo preto e asas vestigiais. Esses dados indicam ligação? Qual é a frequência de recombinação? Represente o cruzamento, mostrando o arranjo dos marcadores genéticos nos cromossomos.
- 7.13** Em coelhos, o alelo dominante C é necessário para pelagem colorida; o alelo recessivo c produz pelagem incolor (albino). Na presença de no mínimo um alelo C , outro gene determina se a pelagem é preta (B , dominante) ou castanha (b , recessivo). Uma linhagem homozigota de coelhos com pelagem castanha foi cruzada com uma linhagem homozigota de albinos. Em seguida, a F_1 foi cruzada com coelhos homozigotos duplos recessivos, produzindo os seguintes resultados: 34 de pelagem preta; 66 de pelagem castanha; 100 de pelagem albina. Há ligação entre os genes b e c ? Qual é a frequência de recombinação? Represente o cruzamento, mostrando o arranjo dos marcadores genéticos nos cromossomos.
- 7.14** Em tomateiros, a planta alta (D) é dominante em relação à planta anã (d) e o fruto esférico (P) é dominante em relação ao fruto piriforme (p). Os genes para altura da planta e formato do fruto são ligados com recombinação de 20% entre eles. Uma planta alta (I) com fruto esférico foi cruzada com uma planta anã de fruto piriforme. O resultado foi: 81 plantas altas com fruto esférico; 79 plantas anãs com fruto piriforme; 22 plantas altas com fruto piriforme; 17 plantas anãs com fruto esférico. Outra planta alta de fruto esférico (II) foi cruzada com uma planta anã de fruto piriforme, com o seguinte resultado: 21 plantas altas com fruto piriforme; 18 plantas anãs com fruto esférico; 5 plantas altas com fruto esférico; 4 plantas anãs com fruto piriforme. Represente esses dois cruzamentos mostrando os marcadores genéticos nos cromossomos. Que classes fenotípicas você esperaria do cruzamento das duas plantas altas com frutos esféricos, isto é, $I \times II$? Em que proporções?
- 7.15** Em *Drosophila*, os genes sr (tórax listrado [*stripe*]) e e (corpo ébano [*ebony*]) estão localizados a 62 e 70 cM, respectivamente, da extremidade esquerda do cromossomo 3. Uma fêmea listrada homozigota para e^+ foi cruzada com um macho de corpo ébano homozigoto para sr^+ . Toda a prole apresentou fenótipo selvagem (corpo cinza e sem listras).
- (a) Que tipos de gametas serão produzidos pelas fêmeas da F_1 ? Em que proporções?
- (b) Que tipos de gametas serão produzidos pelos machos da F_1 ? Em que proporções?
- (c) Qual é a prole esperada do cruzamento de fêmeas da F_1 com machos de corpo ébano e listrado? Em que proporções?
- (d) Qual seria a prole esperada do intercruzamento de machos e fêmeas da F_1 ? Em que proporções?
- 7.16** Em *Drosophila*, os genes a e b estão nas posições 22,0 e 42,0 no cromossomo 2, e os genes c e d estão nas posições 10,0 e 25,0 no cromossomo 3. Uma mosca homozigota para os alelos selvagens desses quatro genes foi cruzada com uma mosca homozigota para os alelos recessivos, e as fêmeas da F_1 foram retrocruzadas com os pais recessivos quádruplos. Qual seria a prole esperada desse retrocruzamento? Em que proporções?
- 7.17** Os genes de *Drosophila* vg (asas vestigiais) e cn (olhos cinabre) estão localizados em 67,0 e 57,0,

respectivamente, no cromossomo 2. Uma fêmea de uma linhagem homozigota de moscas com asas vestigiais foi cruzada com um macho de uma linhagem homozigota de moscas com olhos cinabre. Os híbridos da F₁ tinham fenótipo selvagem (asas longas e olhos vermelho-escuros).

- (a) Quantos tipos diferentes de gametas poderiam produzir as fêmeas da F₁? Em que proporções?
 (b) Que tipo de prole seria esperado do cruzamento dessas fêmeas com machos de olhos cinabre e asas vestigiais? Em que proporções?

7.18 Em *Drosophila*, os genes *st* (olhos escarlate [*scarlet*]), *ss* (ausência de cerdas [*spineless*]) e *e* (corpo ébano [*ebony*]) estão localizados no cromossomo 3, e as posições de mapa são:

<i>st</i>	<i>ss</i>	<i>e</i>
44	58	70

Todas essas mutações são recessivas em relação ao alelo selvagem (*st*⁺, olhos vermelho-escuros; *ss*⁺, cerdas lisas; *e*⁺, corpo cinza). Fêmeas de fenótipo selvagem com genótipo *st ss e*⁺/*st*⁺ *ss*⁺ *e* foram cruzadas com machos recessivos triplos. Preveja os fenótipos da prole e as frequências com que ocorrem supondo que (a) não haja interferência e (b) haja interferência completa.

7.19 No milho, os genes *Pl* para folhas roxas (*purple*) (dominante em relação a *pl*, para folhas verdes), *sm* para estilo-estigma salmão (*salmon*) (recessivo em relação a *Sm*, para estilo-estigma amarelo) e *py* para vegetal pigmeu (*pigmy*) (recessivo em relação a *Py*, para vegetal de tamanho normal) estão no cromossomo 6, com as seguintes posições de mapa:

<i>pl</i>	<i>sm</i>	<i>py</i>
45	55	65

Híbridos do cruzamento *Pl sm py/pl sm py* × *pl Sm Py/pl sm py* foram submetidos a cruzamento-teste com plantas *pl sm py/pl sm py*. Preveja os fenótipos da prole e as frequências com que ocorrem supondo que (a) não haja interferência e (b) haja interferência completa.

7.20 No milho, os genes *Tu*, *j2* e *gl3* estão localizados no cromossomo 4 nas posições de mapa 101, 106 e 112, respectivamente. Se plantas homozigotas para os alelos recessivos desses genes forem cruzadas com plantas homozigotas para os alelos dominantes, e for realizado cruzamento-teste das plantas da F₁ com plantas recessivas triplas, que genótipos você espera encontrar? Em que proporções? Considere que a interferência é completa nesse intervalo de mapa.

7.21 Um geneticista de *Drosophila* fez um cruzamento entre fêmeas homozigotas para as três mutações recessivas ligadas ao X (*y*, corpo amarelo; *ec*, formato de olhos equino; *w*, cor dos olhos branca) e machos de tipo selvagem. Em seguida, cruzou as fêmeas da F₁ com machos mutantes triplos e obteve os seguintes resultados:

Fêmeas	Machos	Número
+ + + / <i>y ec w</i>	+ + +	475
<i>y ec w</i> / <i>y ec w</i>	<i>y ec w</i>	469
<i>y</i> + + / <i>y ec w</i>	<i>y</i> + +	8
+ <i>ec w</i> / <i>y ec w</i>	+ <i>ec w</i>	7
<i>y</i> + <i>w</i> / <i>y ec w</i>	<i>y</i> + <i>w</i>	18
+ <i>ec</i> + / <i>y ec w</i>	+ <i>ec</i> +	23
+ + <i>w</i> / <i>y ec w</i>	+ + <i>w</i>	0
<i>y ec</i> + / <i>y ec w</i>	<i>y ec</i> +	0

Determine a ordem dos três loci *y*, *ec* e *w* e estime as distâncias entre eles no mapa de ligação do cromossomo X.

7.22 Um geneticista de *Drosophila* fez um cruzamento entre fêmeas homozigotas para as três mutações ligadas ao X (*y*, corpo amarelo; *B*, formato de olhos em barra; *v*, cor vermelhão dos olhos) e machos de tipo selvagem. As fêmeas da F₁, que tinham corpo cinza e olhos em barra com pigmento vermelho-escuro, foram cruzadas com machos *y B*⁺ *v*, com o seguinte resultado:

Fenótipo	Número
corpo amarelo, olhos em barra, } vermelhão	546
tipo selvagem	
corpo amarelo } olhos em barra, vermelhão	244
corpo amarelo, olhos vermelhão } olhos em barra	160
corpo amarelo olhos em barra } olhos vermelhão	50

Determine a ordem desses três loci no cromossomo X e estime as distâncias entre eles.

7.23 O cruzamento-teste de fêmeas de *Drosophila* heterozigotas para três mutações recessivas *e* (corpo ébano), *st* (olhos escarlate) e *ss* (ausência de cerdas) obteve a seguinte prole:

Fenótipo	Número
tipo selvagem	67
corpo ébano	8
corpo ébano, olhos escarlate	68
corpo ébano, ausência de cerdas	347
corpo ébano, olhos escarlate, ausência de cerdas	78
olhos escarlate	368
olhos escarlate, ausência de cerdas	10
ausência de cerdas	54

- (a) O que indica que os genes estão ligados?
 (b) Qual era o genótipo das fêmeas heterozigotas originais?
 (c) Qual é a ordem dos genes?
 (d) Qual é a distância de mapa entre *e* e *st*?
 (e) Entre *e* e *ss*?

- (f) Qual é o coeficiente de coincidência?
(g) Faça o diagrama dos cruzamentos desse experimento.

7.24 Considere uma fêmea de *Drosophila* com o seguinte genótipo do cromossomo X:

$$\frac{w \text{ } dor^+}{w^+ \text{ } dor}$$

Os alelos recessivos *w* e *dor* causam cores dos olhos mutantes (branca e laranja-escuro, respectivamente). No entanto, *w* é epistático em relação a *dor*; isto é, os genótipos *w dor/Y* e *w dor/w* produzem olhos brancos. Se houver recombinação de 40% entre *w* e *dor*, que proporção dos filhos machos dessa fêmea heterozigota terá fenótipo mutante? Que proporção terá olhos vermelhos ou laranja-escuro?

7.25 Em *Drosophila*, as mutações recessivas ligadas ao X *prune* (*pn*) e *garnet* (*g*) recombinam-se com frequência de 0,4. Essas duas mutações tornam os olhos castanhos em vez de vermelho-escuros. Fêmeas homozigotas para a mutação *pn* foram cruzadas com machos hemizigotos para a mutação *g*, e as filhas da F₁, todas com olhos vermelho-escuros, foram cruzadas com os irmãos de olhos castanhos. Preveja a frequência de filhos machos que terão olhos vermelho-escuros nesse último cruzamento.

7.26 Suponha que haja em *Drosophila* três genes ligados ao X, *x*, *y*, e *z*, e que cada alelo mutante é recessivo em relação ao alelo selvagem. O cruzamento entre fêmeas heterozigotas para esses três loci e machos de tipo selvagem produziu a seguinte prole:

Fêmeas	+	+	+	1.010
Machos	+	+	+	39
	+	+	z	430
	+	y	z	32
	x	+	+	27
	x	y	+	441
	x	y	z	31
Total: 2.010				

Usando esses dados, construa um mapa de ligação gênica dos três genes e calcule o coeficiente de coincidência.

7.27 No nematódeo *Caenorhabditis elegans*, os genes ligados *dpy* (corpo truncado [*dumpy*]) e *unc* (comportamento descoordenado [*uncoordinated*]) recombinam-se com uma frequência *P*. Se um heterozigoto para repulsão com mutações recessivas nesses genes for autofertilizado, que fração da prole terá corpo truncado e movimentos descoordenados simultaneamente?

7.28 No cruzamento-teste a seguir, os genes *a* e *b* estão distantes 20 cM, e os genes *b* e *c* estão distantes 10 cM: *a* + *c*/+ *b* + × *a b c*/+ *b c*. Se o coeficiente de coincidência é 0,5 nesse intervalo no mapa de

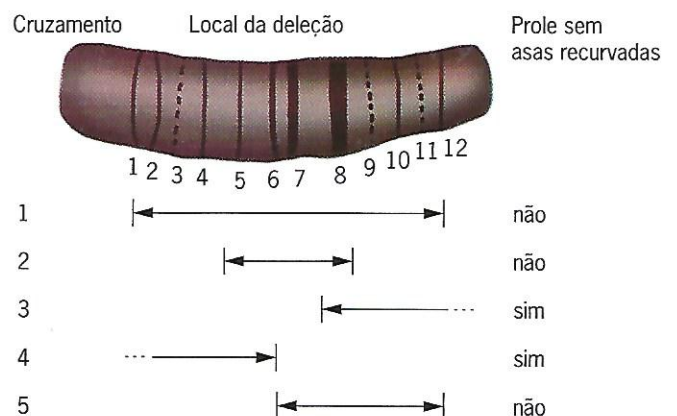
ligação, quantos indivíduos homozigotos recessivos triplos são esperados na prole de 1.000 indivíduos?

7.29 Fêmeas de *Drosophila* heterozigotas para três mutações recessivas, *a*, *b* e *c*, foram cruzadas com machos homozigotos para as três mutações. O resultado do cruzamento foi:

Fenótipo	Número
+ + +	75
+ + <i>c</i>	348
+ <i>b c</i>	96
<i>a</i> + +	110
<i>a b</i> +	306
<i>a b c</i>	65

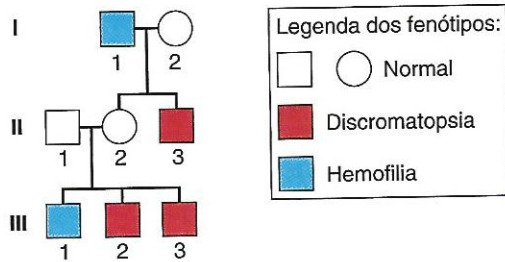
Faça um mapa de ligação mostrando a ordem correta desses três genes e estime as distâncias entre eles.

7.30 Um segundo cromossomo de *Drosophila* que tinha uma mutação letal recessiva, *l(2)g14*, foi mantido em estoque com um cromossomo balanceador marcado por uma mutação dominante para asas recurvadas (*curly*). Essa última mutação, designada *Cy*, também está associada a um efeito letal recessivo, mas esse efeito é diferente do efeito de *l(2)g14*. Assim, as moscas *l(2)g14/Cy* sobrevivem e têm asas recurvadas. Moscas sem a mutação *Cy* têm asas retas. Um pesquisador cruzou fêmeas *l(2)g14/Cy* com machos que tinham os segundos cromossomos com diferentes deleções (todos homozigotos letais) balanceados em relação ao cromossomo *Cy* (genótipo *Df/Cy*). Cada cruzamento foi classificado quanto à presença ou ausência de prole com asas retas.

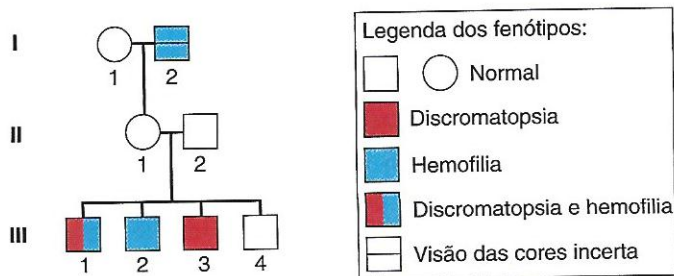


Em que banda está localizada a mutação letal *l(2)g14*?

7.31 O heredograma a seguir, apresentado em 1937 por C. L. Birch, mostra a herança de discromatopsia e hemofilia ligadas ao X em uma família. Qual é o genótipo de II-2? Algum de seus filhos apresenta evidências de recombinação entre os genes para discromatopsia e hemofilia?



7.32 O heredograma a seguir, apresentado em 1938 por B. Rath, mostra a herança de discromatopsia e hemofilia ligadas ao X em uma família. Quais são os possíveis genótipos de II-1? Para cada genótipo possível, avalie se os filhos de II-1 têm evidências de recombinação entre os genes para discromatopsia e hemofilia.



7.33 Uma mulher normal, cujo pai tinha discromatopsia, casou-se com um homem normal, e seu primeiro filho, um menino, é hemofílico. Tanto a discromatopsia quanto a hemofilia são causadas por

mutações recessivas ligadas ao X, e os genes em questão são separados por 10 cM. Esse casal pretende ter outro filho. Qual é a probabilidade de que ele tenha hemofilia? Discromatopsia? Hemofilia e discromatopsia? Nem hemofilia nem discromatopsia?

7.34 Duas linhagens de milho, M1 e M2, são homozigotas para quatro mutações recessivas, *a*, *b*, *c* e *d*, em um dos grandes cromossomos no genoma. A linhagem W1 é homozigota para os alelos dominantes dessas mutações. Os híbridos produzidos pelo cruzamento de M1 e W1 apresentam muitas classes diferentes de recombinantes, enquanto os híbridos produzidos pelo cruzamento de M2 e W1 não têm recombinantes. Qual é a diferença entre M1 e M2?

7.35 Um geneticista estudioso de *Drosophila* identificou uma linhagem de moscas com uma grande inversão no braço esquerdo do cromossomo 3. Essa inversão inclui duas mutações, *e* (corpo ébano [*ebony*]) e *cd* (olhos vermelho-vivo [*cardinal*]), e é ladeada por duas outras mutações, *sr* (tórax listrado [*stripe*]) à direita e *ro* (olhos rugosos [*rough*]) à esquerda. O geneticista quer substituir as mutações *e* e *cd* dentro da inversão por seus alelos selvagens e pretende fazer isso recombinando o cromossomo invertido, mutante múltiplo, com um cromossomo sem inversão, de tipo selvagem. Com que ocorrência o geneticista está contando para alcançar seu objetivo? Explique.

Genômica na Web em <http://www.ncbi.nlm.nih.gov>

Os mapas cromossômicos foram desenvolvidos por T. H. Morgan e seus alunos, que usaram *Drosophila* como organismo experimental.

1. Encontre no mapa genético as posições dos genes *w* (olhos brancos), *m* (asas em miniatura) e *f* (cerdas bifurcadas) no cromossomo X (também designado cromossomo 1) de *Drosophila melanogaster*.
2. Encontre as posições desses três genes no mapa citogenético do cromossomo X de *D. melanogaster*.

Dica: No site, clique em *Genomes and Maps*, depois em *Genome Project* e, por fim, em *Genomic Biology*. Então, em *Genoma resources*, clique em *Insects*. Abra a página sobre *Drosophila melanogaster* e, em *Sequencing Centers*, na barra lateral, clique em *FlyBase*, que é o banco de dados de informações genômicas sobre *Drosophila*. Na página principal de *FlyBase*, procure por um dos três genes para encontrar as localizações genéticas e citológicas.

3. Use a função *Map Viewer* no site para localizar *w*, *m* e *f* no ideograma do cromossomo X.
4. Genes homólogos são genes derivados de um ancestral comum. O gene *SRY* responsável pela determinação do sexo em seres humanos está localizado no cromossomo Y. Um homólogo desse gene, denominado *SOX3*, está localizado no cromossomo X. Encontre esses dois genes nos ideogramas dos cromossomos sexuais humanos. Em que bandas estão localizados?
5. *RBMX* e *RBMX* são outro par de genes homólogos nos cromossomos X e Y humanos. Localize esses dois genes em relação a *SOX3* e *SRY*. À luz da história evolutiva dos cromossomos X e Y, qual poderia ser o responsável pelas posições desses dois pares de genes nos cromossomos sexuais?

Dica: Pesquise usando a função “Find in This View” na página *Map Viewer* do site.

Genética de Bactérias e seus Vírus

PANORAMA

Bactérias multirresistentes | Uma bomba-relógio?

Oscar Peterson, filho de noruegueses que imigraram para a fronteira do Minnesota no fim do século 19, era uma criança feliz. Mas essa infância feliz durou pouco. Logo, sua mãe adoeceu gravemente, com uma tosse incessante, dor torácica e febre alta. Sua mãe tinha tuberculose (TB), uma doença temida, causada pela bactéria *Mycobacterium tuberculosis*. A TB é altamente contagiosa porque o *M. tuberculosis* é transmitido por gotículas em aerossol produzidas quando uma pessoa infectada tosse ou espirra. Muitas vezes era fatal, porque não existia tratamento eficaz na época. Prescrevia-se ar fresco e, por isso, a família Peterson dormia com as janelas abertas, mesmo durante os meses de inverno. Por ser a tuberculose tão contagiosa, as famílias que tinham alguém doente viviam em isolamento quase total. Os amigos tinham medo de visitá-las e contrair a doença. A mãe de Oscar morreu quando ele tinha 14 anos, e sua vida imediatamente mudou. Ele abandonou a escola para cuidar dos irmãos mais novos enquanto o pai trabalhava.

Milhares de famílias fronteiriças como os Petersons lutaram para sobreviver ao flagelo da tuberculose na primeira parte do século 20. Depois, Alexander Fleming descobriu a penicilina, e houve uma revolução no tratamento das doenças bacterianas. Durante as décadas de 1940 e 1950, os cientistas descobriram um arsenal de antibióticos altamente eficazes. Desse modo, houve uma nítida queda da incidência de tuberculose nos EUA durante a década de 1970. Na verdade, muitos médicos acreditavam que a TB poderia ser totalmente erradicada. Infelizmente, estavam errados!

Em 16 de novembro de 1991, uma manchete no *The New York Times* anunciava: "Tuberculose resistente a medicamentos provoca 13 mortes nas prisões de Nova York". Depois, um agente penitenciário em Siracusa morreu vítima da mesma linhagem de *M. tuberculosis* resistente que matara os presidiários e, infe-

- ▶ Vírus e bactérias em genética
- ▶ Genética dos vírus
- ▶ Genética das bactérias
- ▶ Mecanismos de troca genética em bactérias
- ▶ Significado evolutivo da troca genética em bactérias

lizmente, essa linhagem resistente era apenas a ponta do *iceberg*. Hoje, muitas linhagens de *M. tuberculosis* são resistentes a toda uma série de fármacos e antibióticos. Essas linhagens resistentes são de dois tipos: linhagens multirresistentes (MDR), que resistem à maioria dos antibióticos normalmente prescritos, e linhagens extremamente resistentes (XDR), resistentes também aos antibióticos usados no tratamento da TB MDR. As linhagens MDR e XDR de *M. tuberculosis* estão presentes em todo o mundo, com frequências particularmente altas nas prisões, desde Nova York até a Sibéria. A base genética dessa TB multirresistente é analisada adiante neste capítulo (veja Plasmídios e epissomos) e no Capítulo 17.



Mycobacterium tuberculosis, a bactéria causadora da tuberculose em seres humanos.

Qual é a gravidade da ameaça que o surgimento das bactérias MDR e XDR representa para a saúde humana? O Dr. Lee Reichman, um dos grandes especialistas mundiais em tuberculose, definiu a *M. tuberculosis* MDR como uma "bomba-relógio". Em todo o mundo, 2 bilhões de pessoas (15 milhões nos EUA) estão infectadas por *M. tuberculosis* latente. Desses, 8,4 milhões desenvolvem TB ativa e 2 milhões morrem anualmente. Em 18 de março de 2010, a

Organização Mundial da Saúde divulgou que a TB MDR e a TB XDR alcançaram níveis recorde, com 440.000 casos mundiais em 2008. Em algumas regiões do mundo, uma em cada quatro pessoas com TB não é tratada corretamente com esquemas antibióticos de referência; elas estão infectadas pelas linhagens MDR e XDR de *M. tuberculosis*. Talvez devamos tomar medidas para enfrentar a crise da TB MDR e XDR agora – antes que a "bomba" exploda.

Vírus e bactérias em genética

Bactérias e vírus fizeram contribuições importantes para a ciência da genética.

Vivemos em um mundo com incontáveis bactérias e vírus. Algumas bactérias, como a *M. tuberculosis*, são prejudiciais; outras, como as que usamos para produzir iogurte, são úteis. As bactérias têm papel importante nos ecossistemas do planeta. Elas erodem as rochas, capturam energia das substâncias no ambiente, fixam o nitrogênio atmosférico em substâncias que podem ser usadas por outros organismos e decompõem o corpo de organismos mortos. Se as bactérias não desempenhassem essas funções, não existiria a vida da maneira que conhecemos. Esses diminutos organismos possibilitam a sobrevivência de grandes organismos multicelulares como nós.

Os geneticistas começaram a estudar as bactérias e seus vírus em meados do século 20, anos depois da consolidação dos princípios de Mendel e da teoria cromossômica da hereditariedade. Para os primeiros geneticistas bacterianos e virais, esses minúsculos organismos pareciam oferecer a possibilidade de ampliar a análise genética até um nível bioquímico mais profundo – na verdade, até as moléculas que constituem os genes e os cromossomos. Como veremos neste capítulo e nos subsequentes, essa perspectiva empolgante foi concretizada. A análise genética de bactérias e vírus possibilitou que pesquisadores sondassem a natureza química dos genes e seus produtos. Tudo que agora chamamos de biologia molecular teve como base o estudo de bactérias e vírus.

Para um cientista pesquisador, as bactérias e os vírus têm diversas vantagens em comparação com organismos como o milho ou a *Drosophila*. A primeira é que são pequenos, reproduzem-se com rapidez e formam grandes populações em questão de dias. Um experimentador pode criar 10^{10} bactérias em um pequeno tubo de cultura; já 10^{10} *Drosophila* ocupariam uma sala de 4 m × 4 m × 4 m. A segunda é que bactérias e vírus podem crescer em meio de cultura definido bioquimicamente. Como os constituintes do meio de cultura podem ser modificados à vontade, um pesquisador é capaz de

identificar as necessidades químicas do organismo e investigar como ele processa essas substâncias durante seu metabolismo. Também é possível acrescentar fármacos como antibióticos ao meio para destruir as bactérias seletivamente. Esse tipo de tratamento possibilita ao pesquisador identificar linhagens resistentes e sensíveis de uma espécie de bactérias – por exemplo, para verificar se o *M. tuberculosis* cultivado em material colhido de um paciente é resistente a determinado antibiótico. A terceira vantagem é que bactérias e vírus têm estruturas e fisiologia relativamente simples. Portanto, são ideais para o estudo de processos biológicos fundamentais. Por fim, é fácil detectar a variabilidade genética entre esses microrganismos. Ao examinarmos bactérias ou vírus, quase sempre constatamos que têm fenótipos diversos e que essas diferenças são hereditárias. Por exemplo, algumas linhagens de espécies bacterianas podem crescer em um meio definido bioquimicamente cuja única fonte de energia é a lactose, mas outras não. As linhagens incapazes de crescer nesse tipo de meio são mutantes em relação ao metabolismo da lactose. A capacidade de obter linhagens mutantes de bactérias e vírus possibilitou que os geneticistas analisassem fenômenos complexos como o recrutamento de energia, a síntese proteica e a divisão celular em nível molecular.

Os avanços da biologia molecular durante as últimas décadas garantiram muitas informações sobre os genomas de muitos vírus e bactérias. Hoje, conhecemos as sequências nucleotídicas completas dos genomas de uma grande quantidade de vírus e bactérias. Essas sequências oferecem informações detalhadas sobre o controle genético do metabolismo em diversas espécies de micróbios e, principalmente, sobre suas relações evolutivas. Examinaremos algumas dessas informações no Capítulo 15 (veja Genômica comparativa).

Neste capítulo nos concentraremos em alguns vírus e bactérias que tiveram papéis importantes na análise genética. Esses microrganismos incluem a bactéria *Escherichia coli* e dois vírus que a infectam. Iniciaremos nossa investigação com os microrganismos mais simples – os vírus que infectam bactérias como a *E. coli*.

PONTOS ESSENCIAIS

- O tamanho pequeno, o tempo de geração curto e as estruturas simples tornaram bactérias e vírus sistemas-modelo úteis para estudos genéticos
- Muitos conceitos básicos de genética foram deduzidos inicialmente a partir de estudos de bactérias e vírus.

Genética dos vírus

Os vírus só se reproduzem ao infectarem células hospedeiras vivas. Os bacteriófagos são vírus que infectam bactérias. Vários conceitos genéticos importantes foram descobertos graças aos estudos de bacteriófagos.

Os vírus situam-se na fronteira entre seres vivos e não vivos. Considere, por exemplo, um vírus que produz manchas nas folhas do tabaco, distúrbio denominado doença do mosaico do tabaco. O vírus do mosaico do tabaco (TMV) pode ser cristalizado e armazenado durante anos. Nesse estado, não apresenta nenhuma das propriedades normalmente associadas aos sistemas vivos: não se reproduz, não cresce nem se desenvolve; não usa energia e não responde a estímulos ambientais. No entanto, se uma suspensão líquida contendo o TMV for esfregada sobre a folha do tabaco, os vírus em suspensão infectarão as células, se reproduzirão, usarão a energia fornecida pelas células vegetais e responderão a sinais celulares. Sem dúvida, eles apresentam propriedades de sistemas vivos.

Na verdade, é a simplicidade dos vírus que os torna recursos de pesquisa ideais para análise genética. Muitas questões difíceis de responder quando se usavam sistemas eucarióticos mais complexos foram resolvidas pelo uso de vírus. No Capítulo 9, nós discutiremos experimentos que usaram vírus para mostrar que as informações genéticas são armazenadas no DNA e no RNA. Nos Capítulos 10, 11 e 12 discutiremos experimentos com uso de vírus para esclarecer os mecanismos de duplicação, transcrição e tradução do DNA. Neste capítulo, nos concentraremos em vírus que infectam bactérias: discorreremos sobre a organização de seus genomas e os métodos desenvolvidos por geneticistas para analisá-los.

Os vírus que infectam bactérias são denominados **bacteriófagos** (do grego, “comer bactérias”). Entre os muitos bacteriófagos identificados, dois tiveram papéis mais importantes na elucidação de conceitos genéticos. Esses dois vírus infectam a *Escherichia coli*, bacilo presente no cólon. Os bacteriófagos são classificados em dois tipos – virulentos e temperados – de acordo com o estilo de vida nas células infectadas. O bacteriófago T4 (fago T4) é um fago virulento; usa o mecanismo metabólico da célula hospedeira para se multiplicar e destrói a célula durante esse processo. O bacteriófago lambda (λ), um fago temperado, é outro colifago (fago que infecta *E. coli*); esse fago, porém, pode destruir a célula hospedeira, como o fago T4, ou fazer uma associação especial com o hospedeiro e replicar seu genoma junto com o genoma da célula hospedeira a cada duplicação celular.

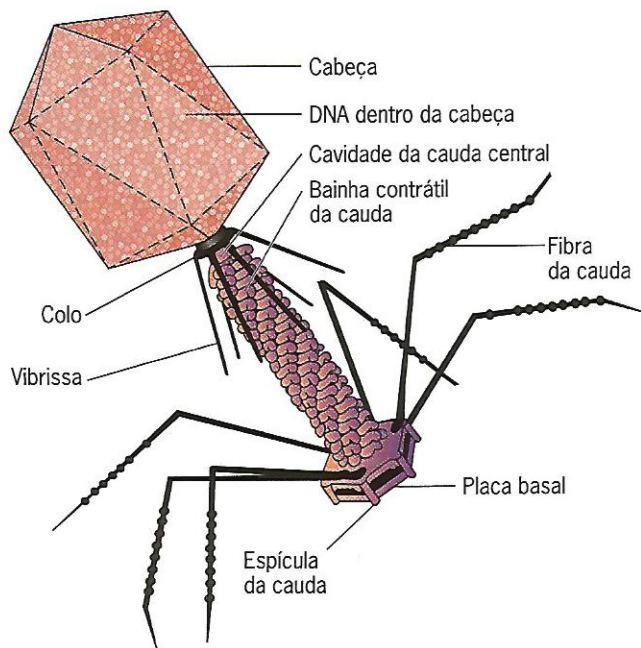
Os resultados de estudos realizados com bacteriófagos T4 e lambda estabeleceram paradigmas genéticos relevantes para compreender outros tipos de vírus, como o vírus da imunodeficiência humana, HIV (veja discussão sobre o HIV no Capítulo 17).

BACTERÍÓFAGO T4

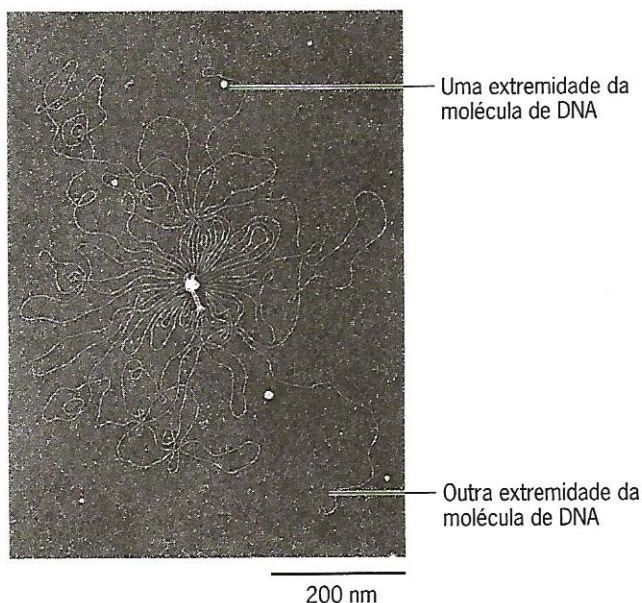
O bacteriófago T4 é um vírus grande cujas informações genéticas são armazenadas em uma molécula de DNA bifilamentar armazenada dentro de uma cabeça proteica (Figura 8.1A). O vírus é quase todo constituído de proteínas e DNA – em quantidades aproximadamente iguais (Figura 8.1B). O cromossomo de T4 tem aproximadamente 168.800 pares de bases e contém cerca de 150 genes caracterizados e um número igual de sequências não caracterizadas que supostamente são genes. A cauda do vírus tem vários componentes importantes. A cavidade central é um canal usado para injetar o DNA do fago na bactéria. A bainha da cauda atua como um pequeno músculo que se contrai e empurra o centro da cauda através da parede da bactéria. As seis fibras da cauda são usadas para localizar receptores na célula hospedeira, e as espículas da cauda na placa basal ligam-se com firmeza a esses receptores. O funcionamento correto de todos esses componentes é imprescindível para que o fago infecte uma célula de *E. coli*.

O bacteriófago T4 é um **fago lítico**; quando infecta uma bactéria, replica-se e destrói a célula hospedeira, produzindo cerca de 300 vírus por célula infectada (Figura 8.2). Depois que é injetado na bactéria hospedeira, o DNA do fago comanda com rapidez (dentro de 2 min) a síntese de proteínas que bloqueiam a transcrição, a tradução e a replicação de genes bacterianos, permitindo que o vírus assuma o controle do mecanismo metabólico do hospedeiro. Alguns genes do fago codificam nucleases que decompõem o DNA do hospedeiro. Outras proteínas do fago iniciam a replicação do DNA do fago. Um pouco mais tarde, são expressos os genes que codificam os componentes estruturais do vírus. Em seguida, começa a montagem dos fagos da prole; a prole infecciosa começa a se acumular na célula hospedeira aproximadamente 17 min após a infecção. Por volta de 25 min depois da infecção, uma enzima codificada pelo fago, denominada lisozima, decompõe a parede celular bacteriana e rompe a bactéria hospedeira, liberando cerca de 300 fagos por célula infectada.

Como já mencionado, o T4 codifica nucleases que degradam o DNA do hospedeiro. Os produtos dessa degra-



A



B

FIGURA 8.1 Bacteriófago T4. **A.** Diagrama mostrando a estrutura do bacteriófago T4. **B.** Micrografia eletrônica de um bacteriófago T4 (centro) cujo DNA foi liberado por choque osmótico. É possível ver as duas extremidades da molécula linear de DNA.

dação são usados na síntese de DNA do fago. Mas como essas enzimas degradam o DNA do hospedeiro sem destruir o DNA do vírus? A resposta é que o DNA de T4 tem uma base incomum – 5-hidroximetilcitosina (HMC; citosina com um grupo – CH_2OH ligado a um dos átomos da molécula de citosina) – em vez de citosina. Além disso, há derivados das moléculas de glicose ligados à HMC. Essas modificações protegem o DNA do T4 contra a decomposição pelas nucleases que decompõem o DNA da célula hospedeira.

É possível mapear os genes nos cromossomos do bacteriófago T4 usando frequências de recombinação, da

mesma maneira que em eucariotos. No entanto, como os vírus têm só um cromossomo que não passa por meiose, o método de mapeamento é um pouco diferente do usado em um organismo como a *Drosophila*. Os cruzamentos são realizados por infecção simultânea de bactérias hospedeiras com dois tipos diferentes de fagos, seguida de pesquisa de genótipos recombinantes na prole de fagos. As distâncias de mapa, em centiMorgans, são calculadas como o número médio de *crossing overs* ocorridos entre marcadores genéticos. Quando as distâncias são curtas, as distâncias de mapa são aproximadamente iguais à porcentagem de cromossomos recombinantes na prole.

Existem muitos tipos diferentes de alelos mutantes no fago T4. As mutações termossensíveis (*ts*) estão entre as mais úteis. O T4 de tipo selvagem cresce em temperaturas que variam de aproximadamente 25°C a 42°C, enquanto os mutantes termossensíveis crescem a 25°C, mas não a 42°C. Assim, é possível distinguir os mutantes *ts* do fago de tipo selvagem por cultura do fago em baixa e alta temperatura. Discutiremos essas mutações termossensíveis e outros tipos de mutantes de T4 nos Capítulos 12 e 13.

BACTERÍÓFAGO LAMBDA

O bacteriófago lambda (λ) é outro colifago que fez grandes contribuições à genética. Ele é menor que o T4; mas seu ciclo de vida é mais complexo. O genoma de lambda contém cerca de 50 genes em uma molécula de DNA bifilamentar com 48.502 pares de bases de comprimento. Essa molécula linear de DNA está na cabeça do λ (Figura 8.3). Logo depois de ser injetado na célula de *E. coli*, a molécula de DNA de λ é convertida em uma forma circular, que participa de todos os processos intracelulares subsequentes.

Dentro da célula, o cromossomo circular de λ pode seguir por duas vias (Figura 8.4). Pode entrar em um ciclo lítico, durante o qual se reproduz e codifica enzimas que lisam a célula hospedeira, assim como o fago T4. Ou pode entrar em uma via lisogênica, na qual é inserido no cromossomo da bactéria hospedeira e, depois, é replicado junto com esse cromossomo. Nesse estado integrado, o cromossomo do λ é denominado **prófago**. Para que esse estado continue, os genes do prófago que codificam produtos participantes da via lítica – por exemplo, enzimas que participam da replicação do DNA do fago, proteínas estruturais necessárias para a morfogênese do fago e a lisozima que catalisa a lise celular – não podem ser expressos.

A integração do cromossomo do λ ocorre por um evento de recombinação sítio-específico entre o DNA circular do λ e o cromossomo circular de *E. coli* (Figura 8.5). Essa recombinação ocorre em sítios de ligação específicos – *attP* no cromossomo de λ e *attB* no cromossomo bacteriano – e é mediada pelo produto do gene *int* do λ , a integrase do λ . A integrase faz a inserção covalente do DNA do λ no cromossomo da célula hospedeira. A recombinação sítio-específica ocorre na região central dos

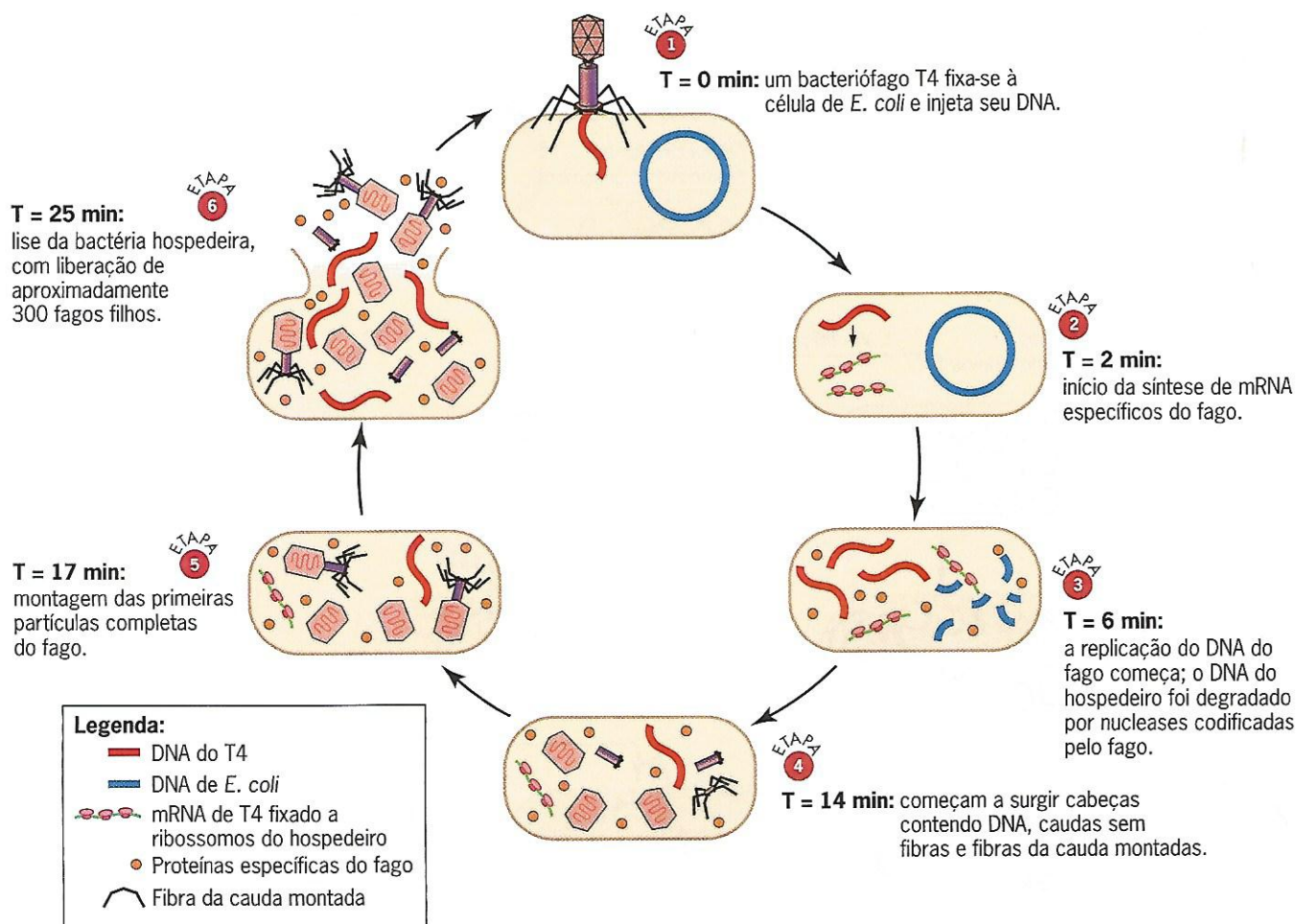


FIGURA 8.2 O ciclo de vida do bacteriófago T4.

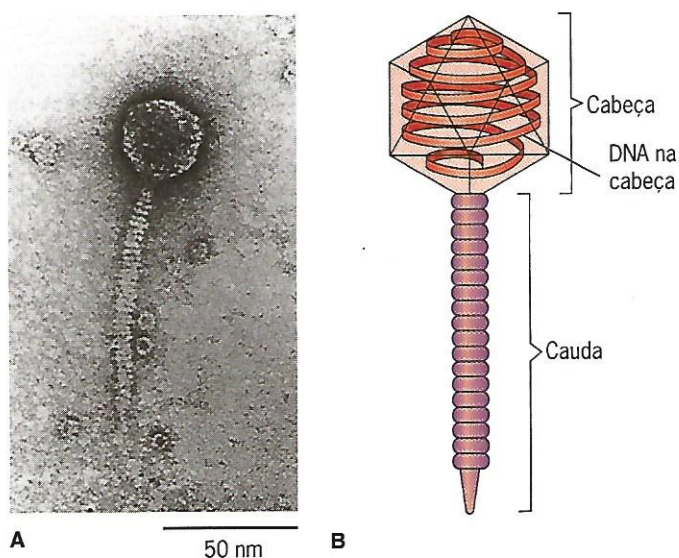
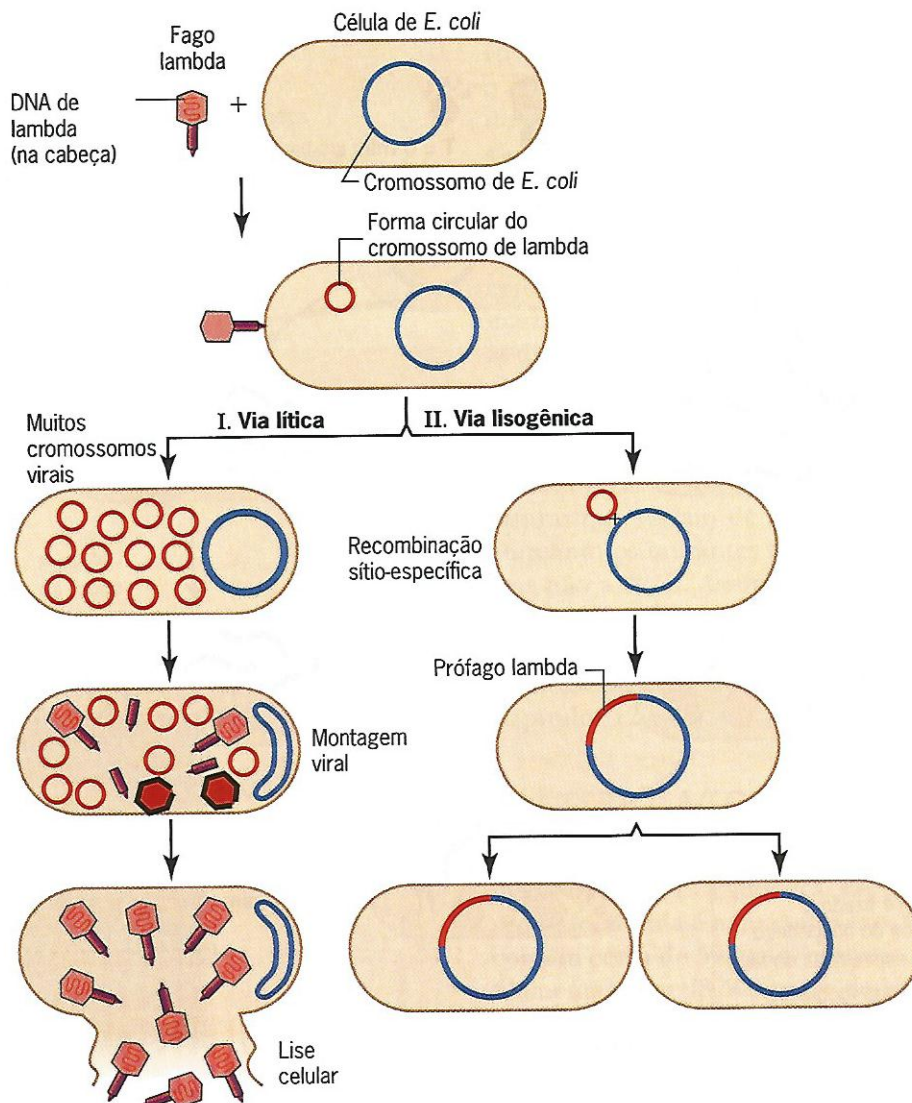


FIGURA 8.3 Bacteriófago λ . Micrografia eletrônica (A) e diagrama (B) mostram a estrutura do bacteriófago λ .



■ FIGURA 8.4 O ciclo de vida do bacteriófago λ . Os dois estados intracelulares do bacteriófago λ : crescimento lítico e lisogenia.

sítios de ligação, onde *attP* e *attB* têm a mesma sequência de 15 pares de nucleotídeos:

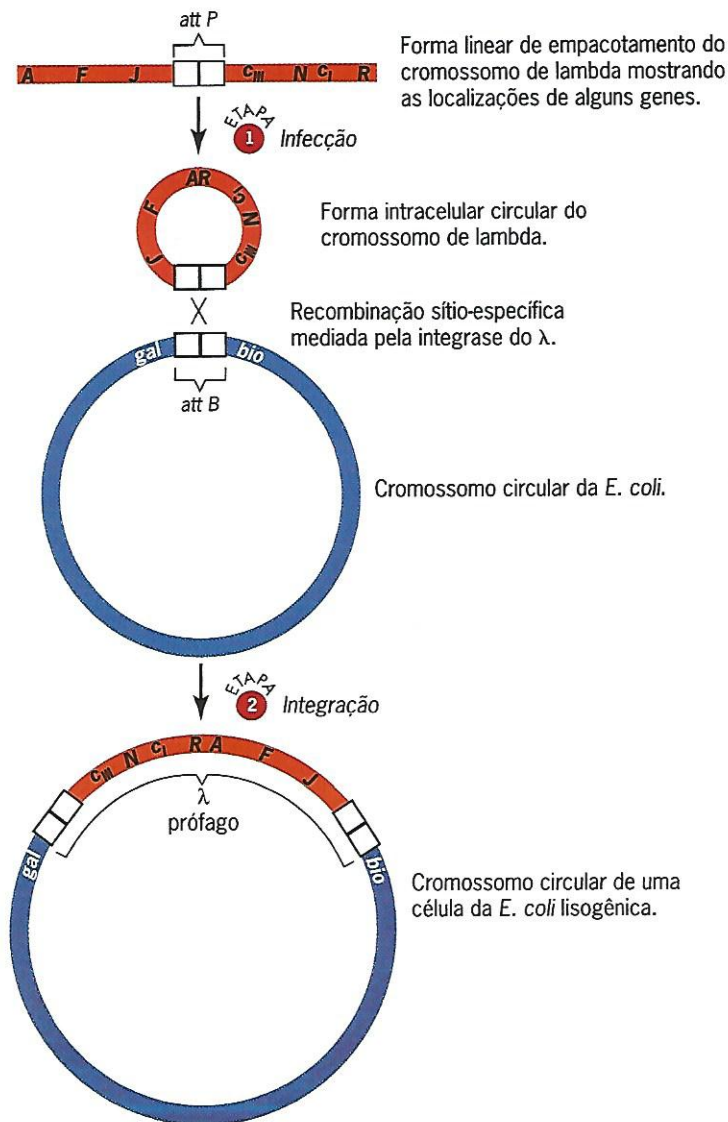
GCTTTTTTATACTAA
CGAAAAATATGATT

Com exceção dessa sequência central, *attP* e *attB* têm sequências muito diferentes. Como a recombinação ocorre nessa sequência central durante a integração, os sítios *attB/P* e *attP/B* resultantes que flanqueiam o prófago integrado também contêm a sequência de 15 pares de nucleotídeos. Essas estruturas são importantes porque facilitam a excisão do prófago por um processo de recombinação sítio-específico muito semelhante.

Cerca de uma vez em cada 10^5 divisões celulares, o prófago λ é excisado espontaneamente do cromossomo do hospedeiro e entra na via lítica. Esse fenômeno explica por que se diz que o prófago está em estado *lisogênico*, isto é, capaz de causar lise, ainda que com baixa frequência. A excisão do prófago λ também pode ser induzida, por exemplo, por irradiação com luz ultravioleta. O processo de excisão geralmente é preciso, com recombinação

sítio-específica entre as sequências centrais em *attB/P* e *attP/B*. Produz um cromossomo do λ autônomo que tem a forma pré-integração original. A excisão requer a integrase do λ e o produto do gene *xis* do λ , a excisase do λ . Essas duas enzimas medeiam um processo de recombinação sítio-específica que é praticamente o inverso do processo de integração. Às vezes, a excisão é anômala, o DNA bacteriano é excisado junto com o DNA do fago. Quando isso ocorre, o vírus resultante é capaz de transferir genes bacterianos de uma bactéria hospedeira para outra. Discutiremos esse processo adiante (veja Mecanismos de troca genética em bactérias).

Estudos com o fago λ contribuíram muito para nosso conhecimento sobre os fenômenos genéticos. Analisaremos a replicação do cromossomo do λ no Capítulo 9. A descoberta do prófago λ (graças à qual André Lwoff foi um dos laureados com o Prêmio Nobel de Fisiologia ou Medicina em 1965) criou o paradigma para os estados provirais do vírus da imunodeficiência humana (HIV) (Capítulo 17) e de vários vírus tumorais de RNA de vertebrados (Capítulo 21).



■ FIGURA 8.5 Integração da molécula de DNA do λ ao cromossomo de E. coli.

PONTOS ESSENCIAIS

- Os vírus são parasitos obrigatórios que só se reproduzem ao infectarem células hospedeiras vivas
- Os bacteriófagos são vírus que infectam bactérias
- O bacteriófago T4 é um fago lítico que infecta E. coli, se reproduz e lisa a célula hospedeira
- O bacteriófago lambda (λ) pode entrar em uma via lítica, como T4, ou em uma via lisogênica, na qual seu cromossomo é inserido no cromossomo da bactéria
- Em seu estado integrado, o cromossomo do λ é denominado prófago, e seus genes líticos mantêm-se inativos.

Genética das bactérias

As bactérias contêm genes que sofrem mutação e produzem fenótipos alterados. A transferência gênica em bactérias é unidirecional – das células doadoras para as receptoras.

As informações genéticas da maioria das bactérias estão armazenadas em um cromossomo principal, que tem

alguns milhares de genes, e em um número variável de “minicromossomos” denominados plasmídios e episomos. Os plasmídios são moléculas de DNA circulares de replicação autônoma que têm de três a várias centenas de genes. Algumas bactérias contêm até 11 diferentes plasmídios além do cromossomo principal. Os episomos são semelhantes aos plasmídios, mas a replicação dos

epissomos pode ser autônoma ou ocorrer como parte do cromossomo principal – em um estado integrado como o prófago λ .

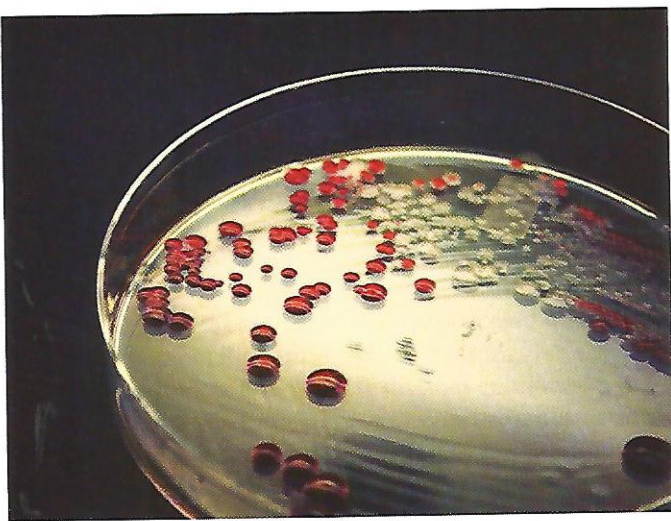
A reprodução das bactérias é assexuada por divisão simples, e cada célula-filha recebe uma cópia do cromossomo. Elas são monoploides, mas “multinucleadas”, ou seja, a célula geralmente contém duas ou mais cópias idênticas do cromossomo. Os cromossomos de bactérias não passam pelos ciclos de condensação mitótica e meiótica que ocorrem durante a divisão celular e a gametogênese em eucariotes. Portanto, os processos de recombinação – distribuição independente e *crossing over* meiótico – que ocorrem durante a reprodução sexuada em eucariotes não ocorrem em bactérias.

Todavia, a recombinação foi tão importante na evolução de bactérias quanto na evolução de eucariotes. Na verdade, processos semelhantes à reprodução sexuada – processos parassexuados – ocorrem em bactérias. Abordaremos esses processos depois de analisar alguns tipos de mutantes usados em genética bacteriana e a natureza unidirecional da transferência de genes entre bactérias.

GENES MUTANTES EM BACTÉRIAS

As bactérias crescem em meio líquido, com frequência exigindo aeração, ou na superfície de um meio semissólido contendo ágar. Se cultivada em meio semissólido, cada bactéria divide-se e cresce de maneira exponencial, produzindo uma colônia visível na superfície do meio. O número de colônias surgidas em uma placa de cultura pode ser usado para estimar o número de bactérias originalmente presentes na suspensão aplicada à placa.

Cada espécie de bactéria produz colônias com cor e morfologia específicas. A *Serratia marcescens*, por exemplo, produz um pigmento vermelho, com formação de colônias vermelhas distintas (Figura 8.6). Mutações nos genes das bactérias podem modificar tanto a cor quanto



■ FIGURA 8.6 Colônias bacterianas. Fotografia mostra colônias da bactéria *Serratia marcescens* que cresce em meio contendo ágar. A cor que distingue as colônias é consequência do pigmento vermelho produzido por essa espécie.

a morfologia da colônia. Além disso, qualquer mutação que reduza a velocidade de multiplicação da bactéria leva à produção de colônias pequenas ou *petites*. Algumas mutações alteram a morfologia da bactéria sem modificar a morfologia da colônia. Além desses mutantes para cor e morfologia da colônia, outros tipos de mutantes foram úteis em estudos genéticos de bactérias.

Mutantes com bloqueio da capacidade de usar fontes energéticas específicas

A *E. coli* de tipo selvagem pode usar praticamente qualquer açúcar como fonte de energia. No entanto, alguns mutantes não usam lactose, o açúcar do leite, para se multiplicar. Eles se multiplicam em outros açúcares, mas não em meio cuja única fonte de energia seja a lactose. Outros mutantes não usam galactose, e ainda outros, a arabinose. A nomenclatura clássica para descrever esses e outros tipos de bactérias mutantes usa abreviaturas de três letras com sobrescritos correspondentes. Nos fenótipos, a primeira letra é maiúscula; nos genótipos, as três letras são minúsculas e em itálico. Portanto, a *E. coli* de tipo selvagem tem fenótipo Lac⁺ (capaz de usar lactose como fonte de energia) e genótipo *lac*⁺. Os mutantes incapazes de usar lactose como fonte de energia têm fenótipo Lac⁻ e genótipo *lac*⁻ (às vezes, apenas *lac*).

Mutantes incapazes de sintetizar um metabólito especial

A *E. coli* de tipo selvagem cresce em meio (meio mínimo) contendo uma fonte de energia e alguns sais inorgânicos. Essas células são capazes de sintetizar todos os metabólitos necessários – aminoácidos, vitaminas, purinas, pirimidinas etc. – a partir dessas substâncias. As bactérias de tipo selvagem são denominadas **prototróficas**. Quando há mutação de um gene que codifica uma enzima necessária para a síntese de um metabólito essencial, a bactéria mutante passa a ter uma nova exigência para se multiplicar. Ela se desenvolve se o metabólito for acrescentado ao meio, mas não se multiplica na ausência dele. Esses mutantes são denominados **auxotróficos**; necessitam de nutrientes auxiliares. Como exemplo, a *E. coli* de tipo selvagem é capaz sintetizar triptofano; essas células têm fenótipo Trp⁺ e genótipo *trp*⁺. Os auxotróficos para triptofano são Trp⁻ e *trp*⁻.

Mutantes resistentes a fármacos e antibióticos

As células de *E. coli* de tipo selvagem são destruídas por antibióticos como ampicilina e tetraciclina. Os fenótipos são Amp^s e Tet^s. Os alelos mutantes que tornam a *E. coli* resistente a esses antibióticos são designados *amp*^r e *tet*^r, respectivamente. As bactérias que têm esses alelos mutantes podem crescer em meio contendo os antibióticos, mas as bactérias de tipo selvagem, não. Assim, os antibióticos podem ser usados para selecionar bactérias que têm genes para resistência. Os genes de resistência atuam como marcadores selecionáveis dominantes.

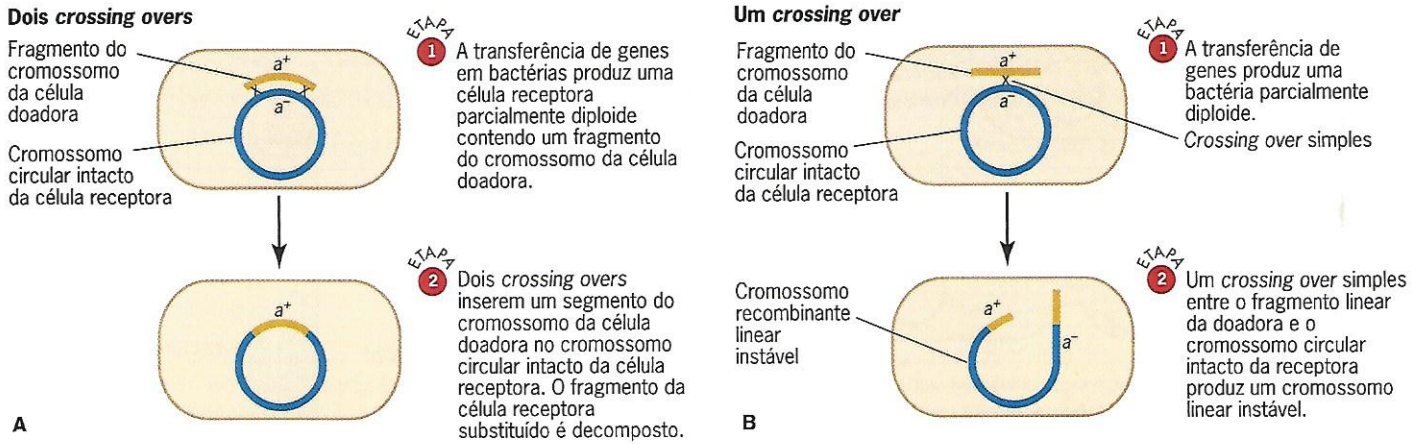


FIGURA 8.7 Recombinação em bactérias. Os processos parassexuais que ocorrem em bactérias produzem diploides parciais que contêm fragmentos lineares do cromossomo das células doadoras e cromossomos circulares intactos das células receptoras. **A.** Para manter a integridade dos cromossomos circulares, os *crossing overs* têm de ocorrer em número par, inserindo segmentos dos cromossomos da célula doadora nos cromossomos da receptora. **B.** Um único *crossing over* entre um fragmento de um cromossomo da célula doadora e um cromossomo circular da receptora destrói a integridade do cromossomo circular, produzindo uma molécula de DNA linear incapaz de se replicar, que, depois, é decomposta.

As bactérias dividem-se rapidamente e produzem grandes populações de células para estudos genéticos. Além disso, é relativamente fácil preparar meios que selecionem genótipos bacterianos específicos (meios seletivos). Logo, as bactérias foram usadas para estudar eventos raros como a mutação de genes e a recombinação entre genes próximos. Comentaremos alguns desses estudos no Capítulo 13.

TRANSFERÊNCIA GÊNICA UNIDIRECIONAL EM BACTÉRIAS

Os processos de recombinação em bactérias implicam a transferência de genes de uma bactéria para outra, e não as trocas recíprocas de genes que ocorrem durante

a meiose em eucariotos. Assim, a transferência de genes é unidirecional, e não bidirecional. A recombinação em bactérias geralmente ocorre entre um fragmento de um cromossomo (da célula doadora) e um cromossomo completo (na célula receptora), e não entre dois cromossomos completos como em eucariotos. Com raras exceções, as células receptoras tornam-se diploides parciais, contendo um trecho linear do cromossomo da doadora e um cromossomo circular completo da receptora. Desse modo, os *crossing overs* têm de ocorrer em número par e inserir um segmento do cromossomo da célula doadora no cromossomo da receptora (Figura 8.7A). Um *crossing over* simples (ou qualquer número ímpar de *crossing overs*) destruirá a integridade do cromossomo da célula receptora, produzindo uma molécula de DNA linear inviável (Figura 8.7B).

PONTOS ESSENCIAIS

- As bactérias geralmente contêm um cromossomo principal
- As bactérias de tipo selvagem são prototróficas; são capazes de sintetizar tudo de que necessitam para se multiplicar e se reproduzir quando têm uma fonte de energia e algumas moléculas inorgânicas
- As bactérias mutantes auxotróficas necessitam de outros metabólitos para seu desenvolvimento
- A transferência gênica em bactérias é unidirecional; genes de uma célula doadora são transferidos para uma célula receptora, sem que haja transferência da receptora para a doadora.

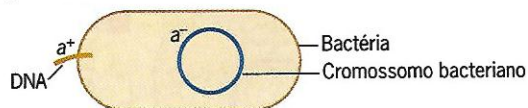
Mecanismos de troca genética em bactérias

As bactérias trocam material genético por três processos parassexuais diferentes.

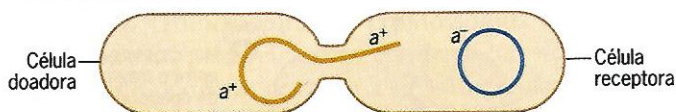
Três processos parassexuais ocorrem em bactérias. A diferença mais óbvia entre esses três processos é o mecanismo de transferência do DNA de uma célula para outra

(Figura 8.8). A **transformação** é a captação de moléculas livres de DNA liberadas de uma bactéria (a célula doadora) por outra bactéria (a célula receptora). A **conjugação** é a transferência direta de DNA de uma célula doadora para uma célula receptora. A **transdução** é a transferência de genes bacterianos de uma célula doadora para uma célula receptora por um bacteriófago.

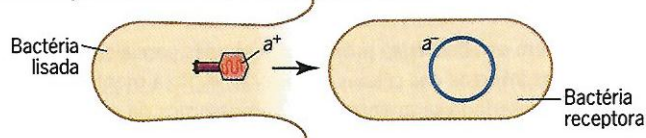
Transformação: captação de DNA livre.



Conjugação: transferência direta de DNA de uma bactéria para outra.



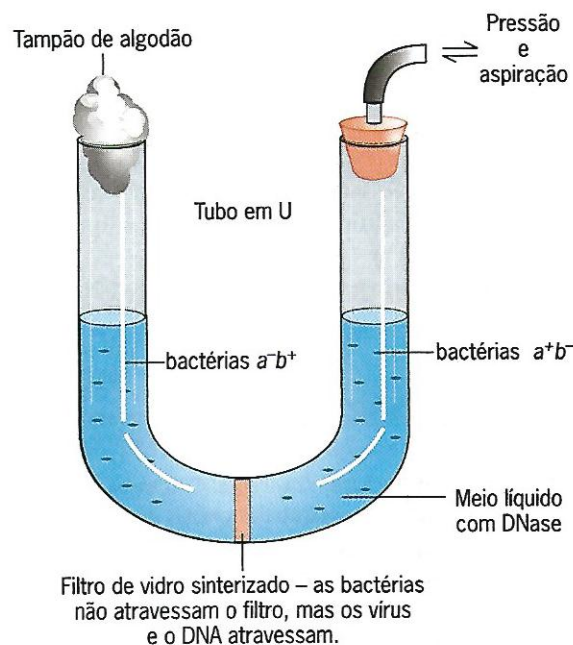
Transdução: transferência de DNA bacteriano por um bacteriófago.



■ FIGURA 8.8 Os três tipos de transferência gênica em bactérias.

Os três processos parassexuais de transferência gênica – transformação, conjugação e transdução – em bactérias podem ser distinguidos por dois critérios simples (Tabela 8.1). (1) O processo requer contato celular? (2) O processo é sensível à desoxirribonuclease (DNase), enzima que degrada DNA? O teste experimental desses dois critérios é muito fácil. A sensibilidade à DNase é determinada pelo simples acréscimo da enzima ao meio de crescimento das bactérias. Se não houver mais transferência de genes, o processo é a transformação. As cápsulas proteicas dos bacteriófagos e as paredes e membranas das bactérias protegem o DNA do doador contra a degradação por DNase durante a transdução e a conjugação, respectivamente.

Um experimento simples pode determinar se o contato celular é ou não necessário para a transferência gênica bacteriana. Nesse experimento, bactérias com diferentes genótipos são postas em braços opostos de um tubo de cultura em U (Figura 8.9). Os dois braços são separados por um filtro de vidro que tem poros suficientemente grandes para permitir a passagem de moléculas de DNA e vírus, mas não de bactérias. Se houver transferência gênica entre as bactérias cultivadas em braços opostos do tubo U, o processo não pode ser a conjugação, que requer contato direto entre células doadoras e receptoras.



■ FIGURA 8.9 O experimento com tubo em U com bactérias. O tubo em U é usado para verificar se a recombinação exige ou não contato celular. Bactérias de diferentes genótipos são postas em cada ramo do tubo, separadas por um filtro de vidro que impede o contato entre elas. Se houver recombinação, o processo não pode ser a conjugação.

Se a transferência gênica observada ocorrer na presença de DNase e na ausência de contato celular, o processo é de transdução.

Os três processos parassexuais não ocorrem em todas as espécies de bactérias; na verdade, a transdução provavelmente é o único que ocorre em todas as bactérias. A ocorrência ou não de transformação ou conjugação em uma espécie depende do surgimento dos genes necessários e do mecanismo metabólico nessa espécie. *E. coli*, por exemplo, não contém genes codificadores das proteínas necessárias para captar o DNA livre. Assim, não há transformação em *E. coli* em condições naturais. Apenas a conjugação e a transdução ocorrem nas células de *E. coli* em habitats naturais. No entanto, cientistas descobriram como transformar células de *E. coli* em laboratório usando tratamentos químicos ou físicos que as tornam permeáveis ao DNA. No Capítulo 14, discutiremos sobre o uso de métodos de transformação para “clonar” (produzir muitas cópias de) genes estranhos em células de *E. coli*.

Tabela 8.1

Distinção entre os três processos parassexuais em bactérias.

Processo de recombinação	Critério	
	Necessidade de contato celular?	Sensível à DNase?
Transformação	não	sim
Conjugação	sim	não
Transdução	não	não

TRANSFORMAÇÃO

Frederick Griffith descobriu a transformação em *Streptococcus pneumoniae* (pneumococo) em 1928. Os pneumococos, como todos os outros seres vivos, exibem variabilidade genética que pode ser reconhecida pela existência de diferentes fenótipos (Tabela 8.2). As duas características fenotípicas importantes na demonstração da transformação por Griffith são (1) presença ou ausência de uma cápsula polissacarídica (polímero de açúcar complexo) ao redor das células bacterianas e (2) o tipo de

Tabela 8.2

Características de linhagens de *Streptococcus pneumoniae* quando cultivadas em meio ágar-sangue.

Tipo	Morfologia da colônia		Cápsula	Virulência	Reação com antissoro preparado contra	
	Aparência	Tamanho			Tipo IIS	Tipo IIIS
IIR ^a	Rugosa	Pequena	Ausente	Avirulento	Nenhuma	Nenhuma
IIS	Lisa	Grande	Presente	Virulento	Aglutinação	Nenhuma
IIIR ^a	Rugosa	Pequena	Ausente	Avirulento	Nenhuma	Nenhuma
IIIS	Lisa	Grande	Presente	Virulento	Nenhuma	Aglutinação

^aEmbora as células tipo R não tenham cápsula, têm genes que orientariam a síntese de um tipo específico (antigênico tipo II ou III) de cápsula se não houvesse bloqueio da formação da cápsula. Quando as células tipo R sofrem retromutação em células tipo S encapsuladas, o tipo de cápsula (II ou III) é determinado por esses genes. Assim, as células R derivadas de células tipo IIS são designadas tipo IIR. Quando essas células tipo IIR sofrem retromutação em células tipo S encapsuladas, as cápsulas são de tipo II.

cápsula, isto é, a composição molecular específica dos polissacarídeos presentes na cápsula. Quando cultivados em ágar-sangue em placas de Petri, pneumococos com cápsulas formam grandes colônias lisas (Figura 8.10) e são designados tipo S (do inglês, *smooth*, liso). Os pneumococos encapsulados são virulentos (patogênicos), causando pneumonia em mamíferos como camundongos e seres humanos. Os pneumococos tipo S virulentos sofrem mutação para uma forma avirulenta (não patogênica), que não tem cápsula polissacarídica, com frequência aproximada de uma em cada 10^7 células. Quando cultivados em meio ágar-sangue, esses pneumococos avirulentos não encapsulados produzem pequenas colônias de superfície rugosa (Figura 8.10) e são, portanto, designados tipo R (do inglês, *rough*, rugoso). A cápsula polissacarídica é necessária para a virulência, porque protege a bactéria contra a destruição por leucócitos. Quando presente, a cápsula pode ser de vários tipos antigênicos diferentes (tipo I, II, III etc.), o que depende da composição molecular específica dos polissacarídeos e, é claro, em última análise, do genótipo da célula.

Os diferentes tipos de cápsula podem ser identificados imunologicamente. A injeção de células tipo II na corrente sanguínea de coelhos leva o sistema imune dos coelhos a produzir anticorpos que reagem especificamente com as células tipo II. Esses anticorpos tipo II aglutinam

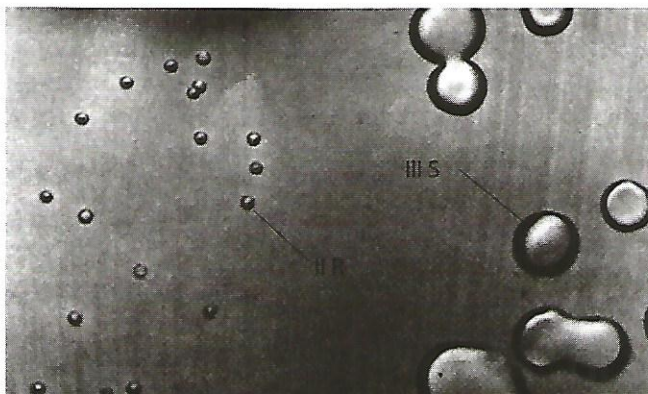


FIGURA 8.10 Fenótipos das colônias das duas linhagens de *Streptococcus pneumoniae* estudadas por Griffith em 1928.

os pneumococos do tipo II, mas não os do tipo I nem do tipo III.

A descoberta inesperada de Griffith foi que, se ele injetasse em camundongos pneumococos tipo IIIS destruídos pelo calor (virulentos quando vivos) mais pneumococos tipo IIR vivos (avirulentos), muitos dos camundongos sucumbiriam à pneumonia, e seriam encontradas células tipo IIIS vivas nos corpos (Figura 8.11). Quando se injetavam apenas pneumococos tipo IIIS destruídos pelo calor,

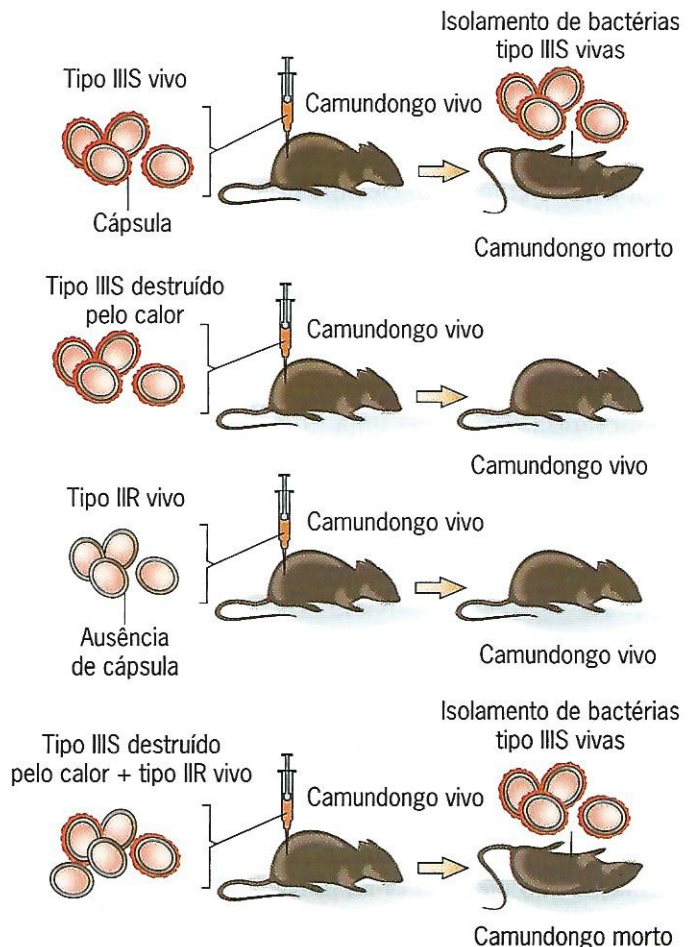


FIGURA 8.11 Descoberta da transformação em *Streptococcus pneumoniae* por Griffith.

nenhum dos camundongos morria. Portanto, a virulência observada não era causada por algumas células tipo IIIS que haviam sobrevivido ao tratamento pelo calor. Os pneumococos patogênicos vivos isolados dos corpos tinham cápsulas polissacarídicas tipo III. Esse resultado é importante porque as células tipo R não encapsuladas podem sofrer mutação de volta em células tipo S encapsuladas. Contudo, quando essa mutação ocorre em uma célula tipo IIR, a célula resultante torna-se IIS, não IIIS. Assim, a transformação de células tipo IIR avirulentas em células tipo IIIS virulentas não é explicável por mutação. Na verdade, alguns componentes das células tipo IIIS mortas (o “princípio transformador”) converteram as células tipo IIR vivas em células tipo IIIS.

Em 1931, experimentos subsequentes realizados por Richard Sia e Martin Dawson mostraram que o fenômeno descrito por Griffith, agora denominado transformação, não era mediado por um hospedeiro vivo. O mesmo fenômeno ocorreu em tubo de ensaio quando se cultivaram células tipo IIR vivas na presença de células tipo IIIS destruídas pelo calor. Como os experimentos de Griffith mostraram que o fenótipo tipo IIIS das células transformadas era transmitido para as células-filhas – isto é, era causado por alteração hereditária permanente no genótipo das células – a demonstração de transformação preparou o terreno para determinar a base química da hereditariedade em pneumococos. Na verdade, a primeira comprovação de que as informações genéticas são armazenadas no DNA, e não nas proteínas, foi a demonstração, em 1944, por Oswald Avery, Colin MacLeod e Maclyn McCarty, de que o DNA era responsável pela transformação em pneumococos. Em vista de seu papel imprescindível no estabelecimento do DNA como o material genético, comentaremos essa demonstração no Capítulo 9.

O mecanismo de transformação foi estudado em muitos detalhes em *S. pneumoniae*, *Bacillus subtilis*, *Haemophilus influenzae* e *Neisseria gonorrhoeae*. O processo básico é semelhante nas quatro espécies; no entanto, há variações de mecanismo em cada uma. *S. pneumoniae* e *B. subtilis* captam DNA de qualquer origem, ao passo que *H. influenzae* e *N. gonorrhoeae* captam apenas seu próprio DNA ou o DNA de espécies bastante próximas. *H. influenzae* e *N. gonorrhoeae* só captam o DNA que contém uma sequência curta especial de pares de nucleotídeos (11 pares de bases em *Haemophilus*; 10 em *Neisseria*) presente em cerca de 600 cópias em seus respectivos genomas.

Mesmo quando a espécie de bactéria é capaz de captar DNA do ambiente, nem todas as células podem fazê-lo. Na verdade, somente as células que expressam os genes codificadores das proteínas necessárias ao processo são capazes de captar DNA. Essas bactérias são ditas **competentes**, e as proteínas que medeiam o processo de transformação são **proteínas de competência (Com)**. As bactérias desenvolvem competência na fase avançada do ciclo de crescimento – quando a densidade celular é alta, mas antes de terminar a divisão celular. O processo pelo qual as células se tornam competentes é mais bem compreendi-

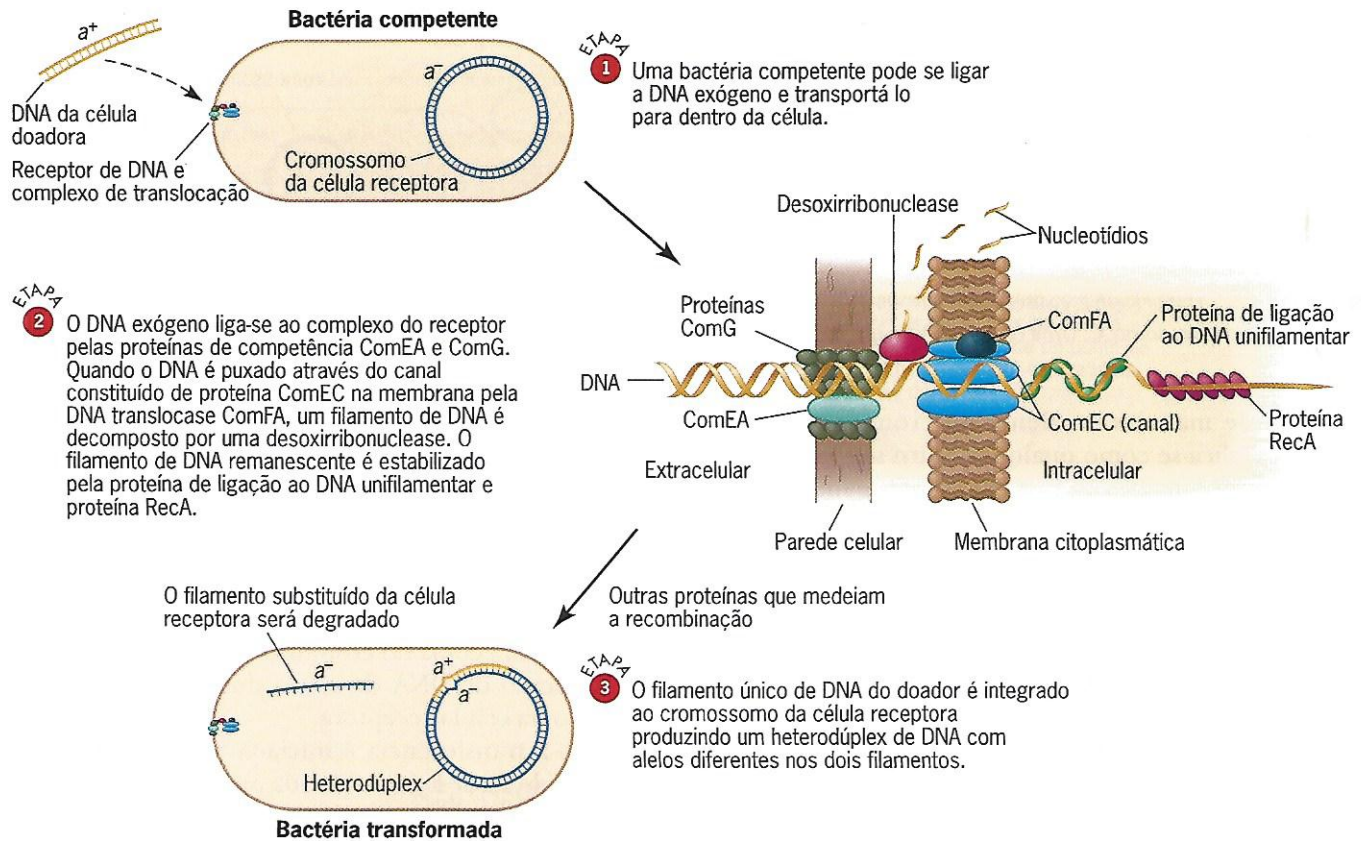
do em *B. subtilis*, em que as células secretam feromônios de competência, pequenos peptídios que se acumulam em alta densidade celular. Altas concentrações dos feromônios induzem a expressão dos genes codificadores de proteínas necessárias à transformação.

Concentremo-nos no mecanismo de transformação em *B. subtilis* (Figura 8.12). Os genes de competência estão localizados em grupos, e cada grupo é designado por uma letra – por exemplo, A, B, C. O primeiro gene em cada grupo é designado A, o segundo, B, e assim por diante. Desse modo, a proteína codificada pelo primeiro gene do quinto grupo é designada ComEA. As proteínas ComEA e ComG ligam o DNA bifilamentar às superfícies de células competentes. Quando o DNA ligado é puxado para o interior da célula pela DNA translocase ComFA (enzima que move ou “transloca” o DNA), um filamento de DNA é decomposto por uma desoxirribonuclease (enzima que degrada o DNA), e o outro filamento é protegido contra a degradação por um revestimento de proteína de ligação ao DNA unifilamentar e proteína RecA (proteína necessária para recombinação). Com o auxílio da RecA e de outras proteínas mediadoras da recombinação, o filamento único de DNA transformador invade o cromossomo da célula receptora, pareando-se com o filamento complementar de DNA e substituindo o filamento equivalente. Em seguida, o filamento substituído da célula receptora é decomposto. Se as células doadora e receptora tiverem alelos diferentes de um gene, a dupla hélice recombinante formada terá um alelo em um filamento e outro alelo no segundo filamento. Uma dupla hélice de DNA desse tipo é denominada **heterodúplex** (uma dupla hélice “heterozigota”); é dividida em dois homodúplex ao se replicar.

As moléculas de DNA captadas por células competentes durante a transformação geralmente correspondem a apenas 0,2 a 0,5% do cromossomo completo. Portanto, exceto se dois genes estiverem muito próximos, nunca estarão na mesma molécula do DNA transformador. Os transformantes duplos para dois genes (p. ex., a^+ e b^+ , usando uma célula doadora a^+b^+ e uma receptora a^-b^-) necessitarão de dois eventos independentes de transformação (captação e integração de uma molécula de DNA com a^+ e de outra molécula com b^+). A probabilidade de que esses dois eventos independentes ocorram juntos é igual ao produto da probabilidade de cada evento isolado. Por outro lado, dois genes próximos podem estar na mesma molécula de DNA transformador, com o surgimento de transformantes duplos com alta frequência. Portanto, pode-se usar a frequência de cotransformação de dois marcadores genéticos para estimar a distância entre eles no cromossomo do hospedeiro.

CONJUGAÇÃO

A transformação não ocorre em *E. coli* – a espécie de bactéria mais estudada – em condições naturais. Assim, poderíamos perguntar se existe algum tipo de transferência gênica entre células de *E. coli*. A resposta a essa



■ **FIGURA 8.12** Mecanismo de transformação em *Bacillus subtilis*. Uma bactéria competente contém um receptor de DNA/complexo de translocação e transporta-o para dentro da célula, onde pode se recombinar ao DNA cromossômico da célula receptora. ComEA, EC, FA e G são proteínas de competência; só são sintetizadas em células competentes. Veja outros detalhes no texto.

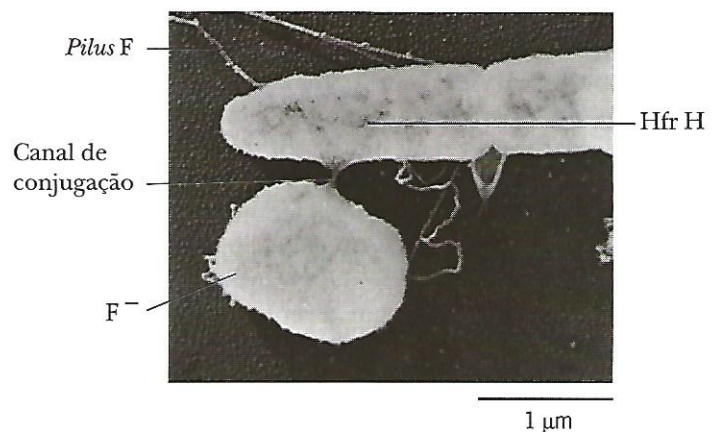
pergunta é “sim”. Em 1946, Joshua Lederberg e Edward Tatum descobriram que as células de *E. coli* transferem genes por conjugação. A conjugação mostrou-se um importante método de mapeamento genético nas espécies de bactérias em que ocorre, e é inestimável em pesquisa genética.

Durante a conjugação, o DNA é transferido de uma célula doadora para uma célula receptora através de um canal de conjugação intracelular especializado que se forma entre elas (**Figura 8.13**). Observe que há contato direto entre as células doadoras e receptoras durante a conjugação; a separação observada na Figura 8.13 foi provocada por forças de estiramento durante o preparo para microscopia.

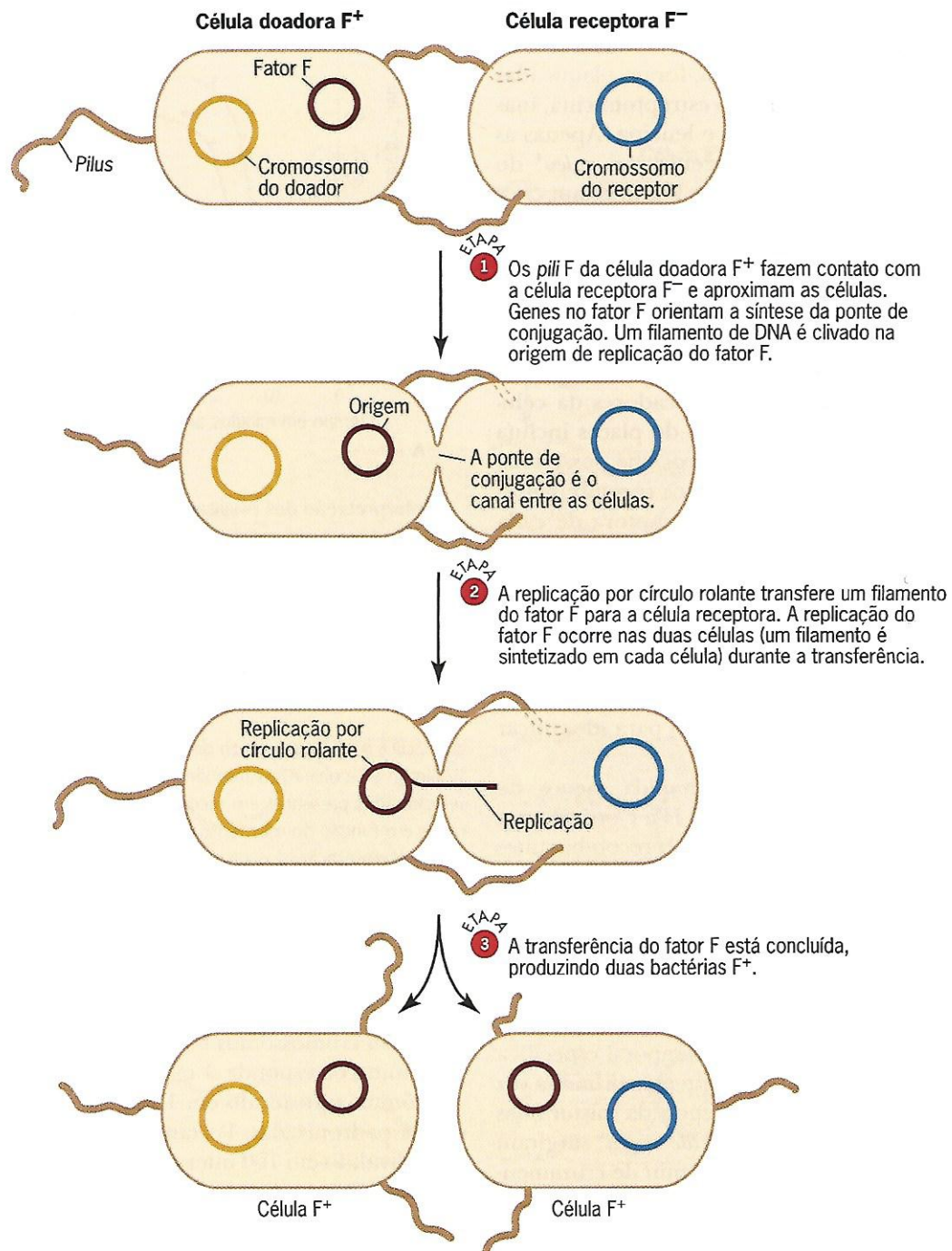
As células doadoras têm apêndices superficiais chamados *pili* F (singular, *pilus* F). A síntese dos *pili* F é controlada por genes presentes em uma pequena molécula circular de DNA chamada **fator F** (*fator de fertilidade*). A maioria dos fatores F tem aproximadamente 10^5 pares de nucleotídios (**Figura 8.20**). As bactérias que contêm um fator F são capazes de transferir genes para outras bactérias. Os *pili* F de uma célula doadora fazem contato com uma célula receptora que não tem fator F e se ligam a essa célula, de maneira que as duas células são postas em contato íntimo. No passado, acreditava-se que o DNA passasse da célula doadora para a célula receptora através de um *pilus* F. Experimentos mais recentes, porém, mostraram que essa ideia é errada. Os *pili* F só participam

do estabelecimento de contato, não da transferência de DNA. Depois que os *pili* aproximam a célula doadora e a célula receptora, forma-se um canal de conjugação entre as células, e, através dele, o DNA é transferido da célula doadora para a célula receptora.

O fator F existe em dois estados: (1) o *estado autônomo*, no qual sua replicação é independente do cromossomo bacteriano, e (2) o *estado integrado*, no qual é



■ **FIGURA 8.13** Conjugação em *E. coli*. Essa antiga micrografia eletrônica, feita por Thomas F. Anderson, mostra a conjugação entre uma célula Hfr H e uma célula F⁻. Na verdade, há justaposição das células doadora e receptora durante a conjugação. O canal de conjugação mostrado foi distendido durante o preparo para microscopia.



■ **FIGURA 8.16** Cruzamento entre células F^+ e F^- . O fator F da célula doadora é replicado durante a transferência de uma célula F^+ para uma célula F^- . Quando o processo termina, cada célula tem uma cópia do fator F.

raros, quando há transferência de um cromossomo Hfr inteiro.

Vários detalhes da conjugação foram decifrados usando uma linhagem específica de Hfr denominada Hfr H (em homenagem a William Hayes, geneticista microbio-inglês, que o isolou). Nessa linhagem, o fator F é integrado perto dos loci *thr* (treonina) e *leu* (leucina), como mostra a Figura 8.15. Em 1957, Elie Wollman e François Jacob, trabalhando no Instituto Pasteur, em Paris, trouxeram uma nova perspectiva ao processo de conjugação por cruzamento de células Hfr H de genótipo *thr⁺ leu⁺ azi^r ton^s lac⁺ gal⁺ str^s* com células F^- de genótipo *thr⁻ leu⁻*

azi^r ton^s lac⁻ gal⁻ str^r. O gene *thr* e o gene *leu* são responsáveis pela síntese dos aminoácidos treonina e leucina, respectivamente. Os pares de alelos *azi^r/azi^s*, *ton^s/ton^r* e *str^s/str^r* controlam a sensibilidade (s) ou a resistência (r) à azida de sódio, ao bacteriófago T1 e à estreptomicina, respectivamente. Os alelos *lac⁺* e *lac⁻* e os alelos *gal⁺* e *gal⁻* determinam a capacidade (+) ou incapacidade (-) de usar lactose e galactose, respectivamente, como fontes de energia.

Em tempos variados depois que as células Hfr H e F^- foram misturadas para iniciar o cruzamento, as amostras foram retiradas e agitadas vigorosamente em agitador

para quebrar as pontes de conjugação e separar as células em conjugação. Essas células, cujo cruzamento havia sido interrompido com tanta indelicadeza, foram plaqueadas em meio que continha o antibiótico estreptomicina, mas não tinha os aminoácidos treonina e leucina. Apenas as células recombinantes que têm os genes *thr*⁺ e *leu*⁺ do genitor Hfr H e o gene *str*^r do genitor F⁻ poderiam crescer nesse meio seletivo. As células doadoras Hfr H seriam destruídas pela estreptomicina, e as células receptoras F⁻ não cresceriam sem treonina e leucina.

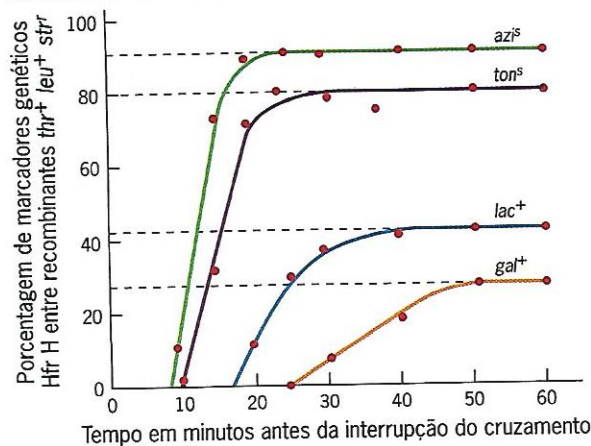
As colônias produzidas pelos recombinantes *thr*⁺ *leu*⁺ *str*^r foram transferidas (Capítulo 13, Figura 13.15) para uma série de placas contendo diferentes meios seletivos para determinar quais dos outros marcadores da célula doadora estavam presentes. A série de placas incluía meios contendo suplementos específicos que possibilitaram a Wollman e Jacob determinar se os recombinantes tinham alelos da célula doadora ou receptora de cada gene. O meio contendo azida de sódio foi usado para distinguir células *azi*^s e *azi*^r. O meio contendo o bacteriófago T1 foi usado para classificar bactérias recombinantes como *ton*^s ou *ton*^r. O meio contendo lactose como única fonte de carbono foi usado para determinar se os recombinantes eram *lac*⁺ ou *lac*⁻, e o meio contendo galactose como única fonte de carbono foi usado para identificar recombinantes *gal*⁺ e *gal*⁻.

Quando a conjugação foi interrompida menos de 8 min depois da mistura das células Hfr H e F⁻, não foram detectados recombinantes *thr*⁺ *leu*⁺ *str*^r. Os recombinantes (*thr*⁺ *leu*⁺ *str*^r) surgiram cerca de 8 min e meio depois da mistura de células Hfr H e F⁻ e acumularam-se até alcançar uma frequência máxima em alguns minutos. Quando se analisou a presença dos marcadores da célula doadora em intervalos variados depois da mistura de células doadoras e receptoras, os alelos das doadoras foram transferidos para células receptoras em uma sequência temporal específica (Figura 8.17). O gene *azi*^s de Hfr H surgiu pela primeira vez em recombinantes cerca de 9 min depois da mistura das bactérias Hfr e F⁻. Os marcadores *ton*^s, *lac*⁺ e *gal*⁺ surgiram pela primeira vez depois de 11, 18 e 25 min de cruzamento, respectivamente. Esses resultados indicaram que os genes de Hfr H estavam sendo transferidos para as células F⁻ em uma ordem temporal específica, refletindo a ordem dos genes no cromossomo (Figura 8.18).

Estudos subsequentes com diferentes linhagens de Hfr mostraram que a transferência gênica poderia ser iniciada em diferentes sítios no cromossomo. Agora sabemos que o fator F pode integrar-se a muitos locais diferentes no cromossomo da *E. coli* e que o local de integração determina onde é iniciada a transferência gênica em cada linhagem de Hfr. Além disso, a orientação da integração do fator F⁻ – seja *d c b a* em sentido horário, seja *a b c d* em sentido horário (Figura 8.15) – determina se a transferência de genes ocorre em sentido horário em relação ao mapa de ligação de *E. coli* ou em sentido anti-horário (Figura 8.19).

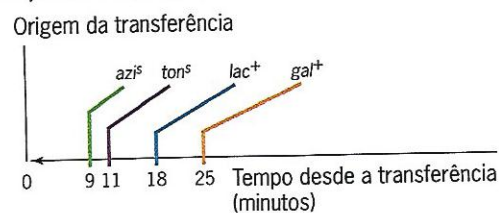
A transferência de um cromossomo completo de uma célula Hfr para uma célula F⁻ leva cerca de 100 min, e a velocidade da transferência parece ser razoavelmente constante. Assim, o tempo necessário para a transferência de

Resumo dos resultados



A

Interpretação dos resultados



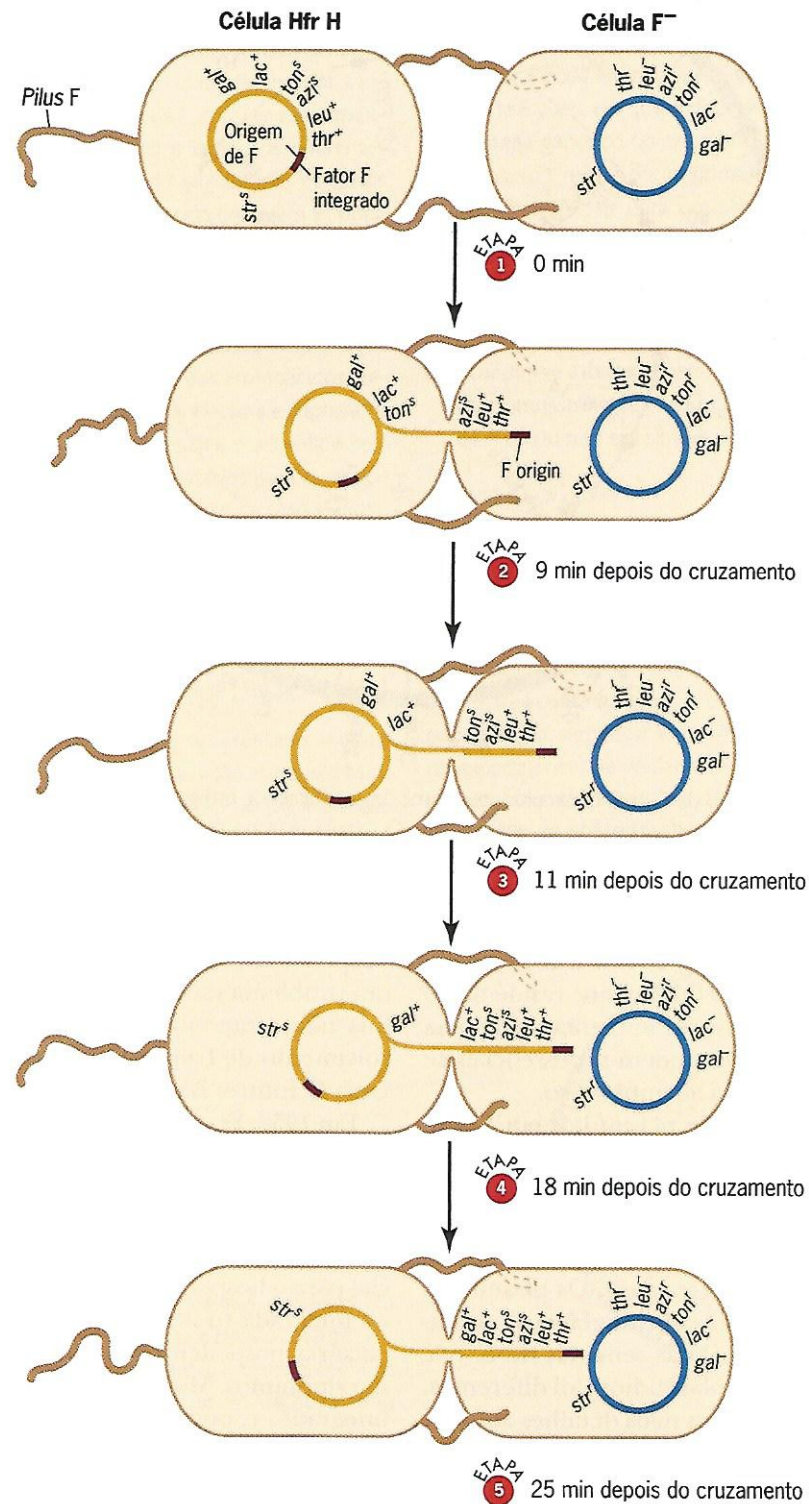
B

■ FIGURA 8.17 Experimento de cruzamento interrompido clássico de Wollman e Jacobs. A. As frequências dos alelos da célula doadora não selecionados presentes em recombinantes *thr*⁺ *leu*⁺ *str*^r são apresentadas em função do momento em que foi interrompido o cruzamento. B. Interpretação dos resultados com base na transferência linear de genes da célula Hfr para a célula F⁻. A transferência é iniciada na origem no fator F, e o momento em que um gene é transferido para a célula F⁻ depende da distância entre ele e o fator F.

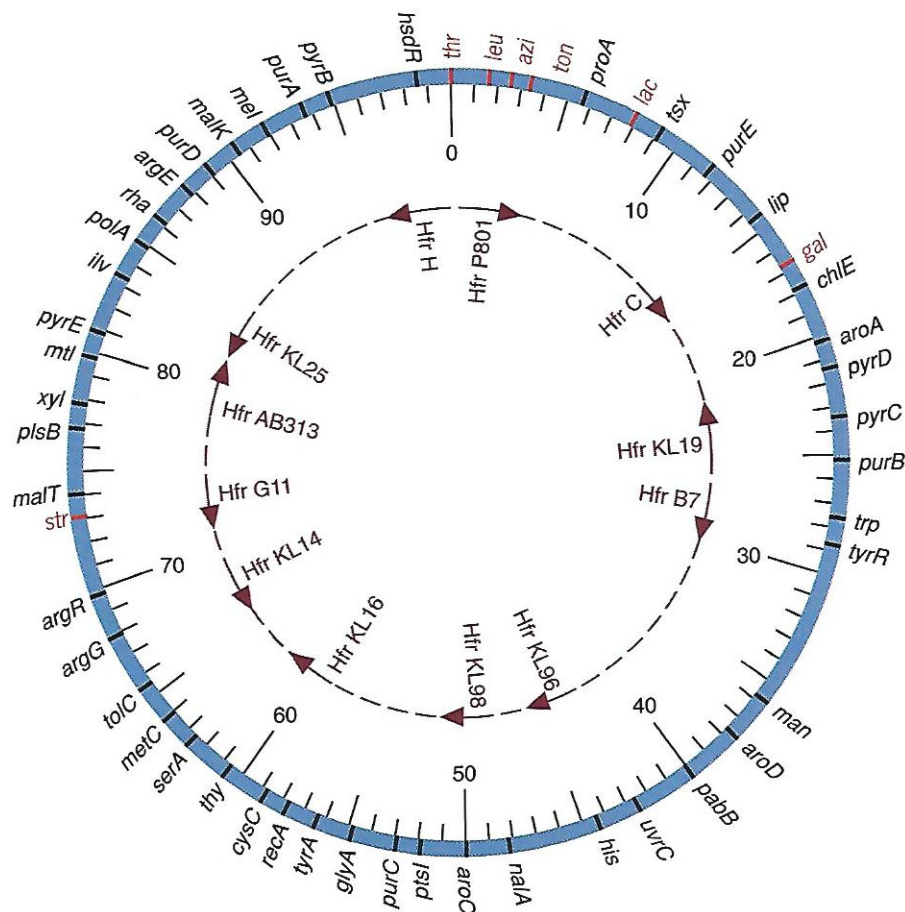
genes durante a conjugação pode ser usado para mapear genes em cromossomos bacterianos. A distância de mapa de 1 min corresponde à extensão de um segmento cromossômico transferido em 1 min de conjugação em condições padronizadas. Portanto, o mapa de ligação de *E. coli* é dividido em 100 intervalos de 1 min (Figura 8.19). A coordenada zero desse mapa circular foi arbitrariamente definida no gene *thrA*. Quando se identifica uma nova mutação em *E. coli*, sua localização no cromossomo é determinada primeiro por mapeamento de conjugação. Para testar seu conhecimento sobre o mapeamento da conjugação, deduza as localizações cromossômicas dos genes analisados no Problema resolvido: Mapeamento de genes com o auxílio de dados de conjugação, no fim desta seção. Em seguida, pode-se usar a transformação ou a transdução para fazer o mapeamento mais preciso.

PLASMÍDIOS E EPISSOMOS

Como já citado, o material genético de uma bactéria está em um cromossomo principal e em uma a várias moléculas de DNA extracromossômico chamadas plasmídios. Por definição, um **plasmídio** é um elemento genético com capacidade de replicação independente do cromossomo principal em um estado extracromossômico. A maioria dos



■ **FIGURA 8.18** Interpretação do experimento de cruzamento interrompido de Wollman e Jacobs. Há transferência linear de genes da célula doadora (Hfr H) para a célula receptora (F⁻). A transferência começa na origem da replicação no fator F integrado e prossegue com a transferência sequencial de genes de acordo com sua localização no cromossomo. O cromossomo replica-se durante o processo de transferência, de maneira que as células Hfr e F⁻ terminam com uma cópia do DNA transferido.



■ **FIGURA 8.19** Mapa circular de ligação da *E. coli*. O círculo interno mostra os locais de integração do fator F em linhagens Hfr selecionadas. As setas indicam se a transferência pela Hfr é horária ou anti-horária. O círculo externo mostra a posição de genes selecionados. O mapa é dividido em 100 unidades, em que cada unidade é o comprimento de DNA transferido durante um minuto de conjugação. Os genes mostrados em vermelho foram usados no famoso experimento de cruzamento interrompido de Wollman e Jacobs (Figuras 8.17 e 8.18).

plasmídios é dispensável ao hospedeiro; ou seja, não é necessária para a sobrevivência da célula em que residem. No entanto, em determinadas condições ambientais, como na presença de um antibiótico, eles podem ser essenciais se tiverem um gene para resistência ao antibiótico.

Existem três tipos principais de plasmídios em *E. coli*: fatores F, plasmídios R e plasmídios Col. Os fatores de fertilidade (F) já foram apresentados (veja Conjugação). Os plasmídios R (plasmídios de resistência) têm genes que tornam as células hospedeiras resistentes aos antibióticos e a outros fármacos antibacterianos. Os plasmídios Col (antes denominados fatores colicinogênicos) codificam proteínas que destroem células sensíveis de *E. coli*. Há uma grande quantidade de plasmídios Col diferentes, mas eles não serão discutidos com mais detalhes aqui.

Alguns plasmídios dotam as células hospedeiras da capacidade de conjugação. Todos os plasmídios F⁺, muitos plasmídios R e alguns plasmídios Col têm essa propriedade; dizemos que são plasmídios conjugativos. Outros plasmídios R e Col não conferem às células a capacidade de conjugação; dizemos que são não conjugativos. A natureza conjugativa de muitos plasmídios R tem um papel importante na rápida disseminação de genes de resistência a antibióticos e fármacos nas populações de bactérias patogênicas, como foi comentado no início deste capítulo. A evolução de plasmídios R que tornam as bactérias

hospedeiras resistentes a múltiplos antibióticos tornou-se um problema médico grave, e o uso de antibióticos para fins não terapêuticos contribuiu para o rápido desenvolvimento de bactérias resistentes a múltiplos fármacos (veja O futuro: Bactérias resistentes a antibióticos).

Em 1958, François Jacob e Elie Wollman reconheceram que o fator F e alguns outros elementos genéticos tinham propriedades especiais. Eles definiram essa classe de elementos e os denominaram epissomos. Segundo Jacob e Wollman, um **epissomo** é um elemento genético não essencial para o hospedeiro, cuja replicação pode ser autônoma ou integrada (por inserção covalente) ao cromossomo da bactéria hospedeira. Os termos *plasmídio* e *epissomo* não são sinônimos. Muitos plasmídios não existem em estados integrados e, portanto, não são epissomos. Da mesma maneira, muitos cromossomos de fagos lisogênicos, como o genoma do fago λ , são epissomos, mas não plasmídios.

A capacidade dos epissomos de se inserirem nos cromossomos depende da presença de sequências curtas de DNA chamadas sequências de inserção (ou elementos IS [*insertion sequences*]). Os elementos IS estão presentes tanto nos epissomos quanto nos cromossomos bacterianos. Essas sequências curtas (cujo comprimento varia de cerca de 800 a 1.400 pares de nucleotídeos) são transponíveis; isto é, podem passar de um cromossomo para outro (Capítulo 17). Além disso, os elementos IS medeiam

PROBLEMA RESOLVIDO



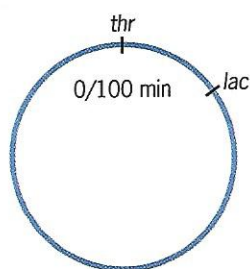
Mapeamento de genes com o auxílio de dados de conjugação

PROBLEMA

Você identificou uma linhagem mutante de *E. coli* que não sintetiza o aminoácido triptofano (Trp^-). Com o objetivo de determinar a localização da mutação trp^- no cromossomo da *E. coli*, fez experimentos de cruzamento interrompido com quatro linhagens diferentes de Hfr. Em todos os casos, as linhagens de Hfr tinham os alelos selvagens dominantes dos genes marcadores, e a linhagem F^- tinha os alelos mutantes recessivos desses genes. O diagrama a seguir mostra o momento de entrada em minutos (entre parênteses) dos alelos selvagens dos genes marcadores na linhagem Trp^- de F^- . Os genes marcadores são thr^+ , aro^+ , his^+ , tyr^+ , met^+ , arg^+ e ilv^+ (que codificam enzimas necessárias para a síntese dos aminoácidos treonina, os aminoácidos aromáticos fenilalanina, tirosina e triptofano, histidina, tirosina, metionina, arginina e isoleucina mais valina, respectivamente) e man^+ , gal^+ , lac^- e xyl^+ (necessários para a capacidade de catabolizar os açúcares manose, galactose, lactose e xilose, respectivamente, e usá-los como fontes de energia).

Hfr A — man^+ (1) trp^+ (9) aro^+ (17) gal^+ (20) lac^+ (29) thr^+ (37)
 Hfr B — trp^+ (6) man^+ (14) his^+ (22) tyr^+ (34) met^+ (42) arg^+ (48)
 Hfr C — thr^+ (3) ilv^+ (20) xyl^+ (25) arg^+ (33) met^+ (39) tyr^+ (47)
 Hfr D — met^+ (2) arg^+ (8) xyl^+ (16) ilv^+ (21) thr^+ (38) lac^+ (46)

No mapa do cromossomo da *E. coli* circular apresentado, indique (1) a localização relativa de cada gene, (2) a posição em que o fator F é integrado a cada uma das quatro células Hfr e (3) a direção da transferência de cromossomo para cada Hfr (sentido horário ou anti-horário; indique a direção com uma seta).



FATOS E CONCEITOS

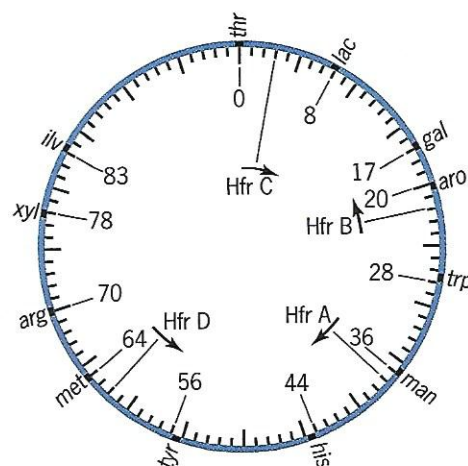
1. O cromossomo da *E. coli* contém uma molécula circular de DNA.
2. O DNA cromossômico é transferido de células doadoras Hfr para células receptoras F^- por replicação por círculo rolante.
3. A replicação por círculo rolante e, portanto, a transferência de genes cromossômicos, inicia-se na origem de replicação no fator F integrado.

4. A direção da transferência (horária ou anti-horária) depende da orientação do fator F no cromossomo de Hfr.
5. O fator F pode integrar-se a muitos locais diferentes no cromossomo da *E. coli* e em qualquer orientação (horária ou anti-horária).
6. O mapa genético do cromossomo da *E. coli* é dividido em minutos, e 1 min é o comprimento do DNA transferido de uma linhagem de Hfr para uma linhagem de F^- durante 1 min de conjugação.
7. A transferência de todo o cromossomo de uma célula Hfr para uma célula F^- leva 100 min; portanto, o mapa de ligação gênica do cromossomo circular completo tem 100 min.
8. Atribuiu-se arbitrariamente ao locus thr a posição "0" no mapa do cromossomo de *E. coli*, com aumento da distância de ligação de 0 a 100 min em sentido horário a partir de thr .

ANÁLISE E SOLUÇÃO

Se examinarmos a sequência de transferência dos genes de cada linhagem de Hfr para a linhagem de F^- , observaremos uma sequência linear em todos os casos.

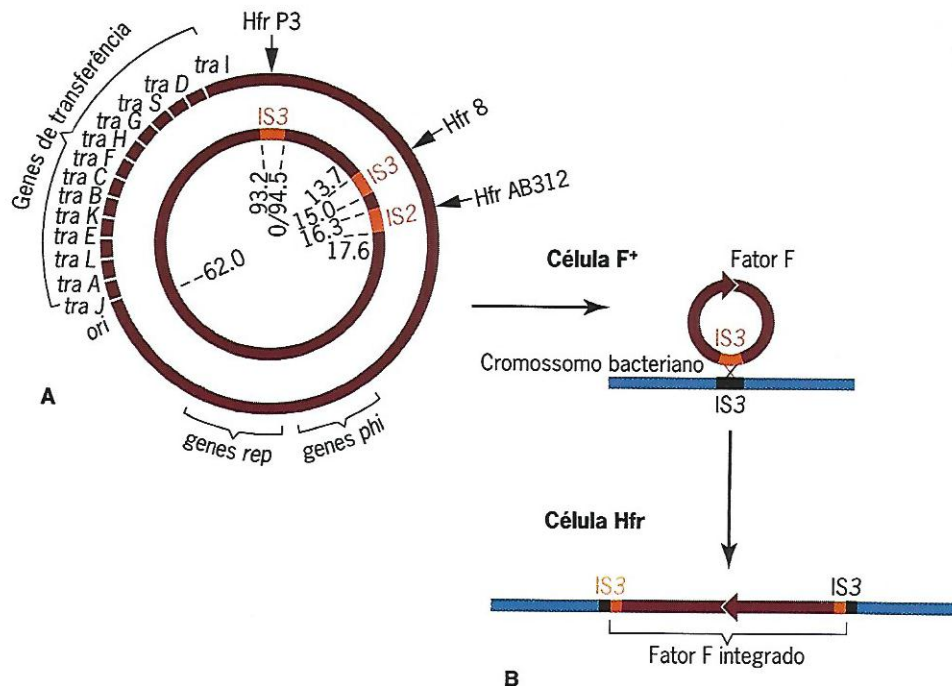
Observe também que seja qual for a sequência de transferência dos genes por diferentes linhagens de Hfr, a distância entre genes adjacentes continua igual. A distância entre man e trp é de 8 min, por exemplo, sem levar em conta o uso da linhagem A ou B de Hfr no experimento. Na verdade, se combinarmos os resultados obtidos usando as quatro linhagens de Hfr e pusermos thr na posição 0, os dados produzirão o mapa genético circular a seguir. O mapa circular é um resultado satisfatório já que sabemos que o DNA cromossômico da *E. coli* também é circular.



a recombinação entre elementos genéticos não homólogos no tocante aos demais aspectos. O papel dos elementos IS na mediação da integração de epissemos é bem documentado no caso do fator F em *E. coli*. O *crossing over* entre elementos IS no fator F e o cromossomo bacteriano produz Hfr com diferentes origens e direções de transferência durante a conjugação (Figura 8.20).

FATORES F' E SEXODUÇÃO

Como discutido na seção anterior, uma linhagem Hfr é produzida pela integração de um fator F ao cromossomo por recombinação entre elementos IS no cromossomo e elementos IS no fator F (Figura 8.20). Você acredita que esse processo de recombinação possa ser reversível?



■ **FIGURA 8.20** Elementos IS medeiam a integração do fator F. **A.** Mapa abreviado da estrutura do fator F em linhagem K12 de *E. coli*, com distâncias em quilobases (1.000 pares de nucleotídeos). As localizações dos genes necessários para transferência conjugada (genes *tra*), replicação (genes *rep*) e inibição do crescimento do fago (genes *phi*) são mostradas junto com as posições de três elementos IS. As setas indicam o elemento IS específico que mediu a integração do fator F durante a formação das linhagens Hfr indicadas. **B.** A recombinação entre elementos IS insere o fator F no cromossomo bacteriano, produzindo uma célula Hfr.

Na verdade, há raras células F⁺ em culturas de Hfr, indicando que há excisão do fator F (por um processo que é basicamente o inverso do evento de integração mostrado na Figura 8.20B). Além disso, eventos de excisão anômala como o que mostra a Figura 8.21 produzem fatores F autônomos que têm genes bacterianos. Esses fatores F modificados, denominados F' ("F linha"), foram identificados pela primeira vez por Edward Adelberg e Sarah Burns em 1959. O tamanho dos fatores F' varia de um gene bacteriano até metade do cromossomo bacteriano (Figura 8.22).

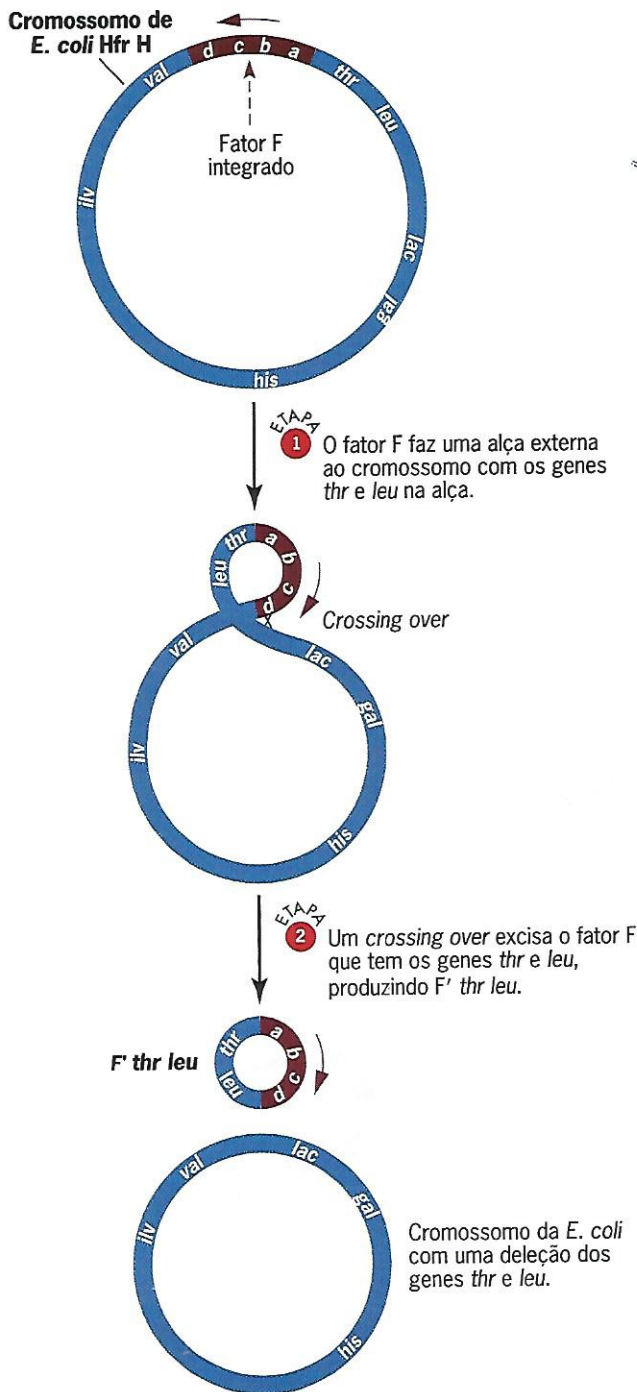
A transferência de fatores F' para células receptoras (F⁻) é denominada **sexdução**; ocorre pelo mesmo mecanismo que a transferência de fator F em cruzamentos F⁺ × F⁻ (Figura 8.16), com uma diferença importante: genes bacterianos incorporados aos fatores F' são transferidos para células receptoras em frequência muito maior. Os fatores F' são instrumentos úteis para estudos genéticos; podem ser usados para produzir diploides parciais com duas cópias de qualquer gene ou conjunto de genes ligados. Portanto, a sexdução pode ser usada para determinar as relações de dominância entre alelos e fazer outros testes genéticos que exigem duas cópias de um gene na mesma célula.

Considere um fator F' *thr*⁺ *leu*⁺ gerado por excisão anômala do fator F de Hfr H, como mostra a Figura 8.21. Os cruzamentos entre células doadoras F' *thr*⁺ *leu*⁺ e células receptoras *thr*⁻ *leu*⁻ produzem diploides parciais *thr*⁻ *leu*⁻/F' *thr*⁺ *leu*⁺. Esses diploides parciais são instáveis porque o fator F' pode ser perdido, produzindo haploides *thr*⁻ *leu*⁻, ou pode haver recombinação entre

o cromossomo e o F', produzindo recombinantes *thr*⁺ *leu*⁺ estáveis. Para analisar com mais detalhes o uso de diploides parciais em mapeamento genético, leia Resolução! Como mapear genes próximos usando diploides parciais?

TRANSDUÇÃO

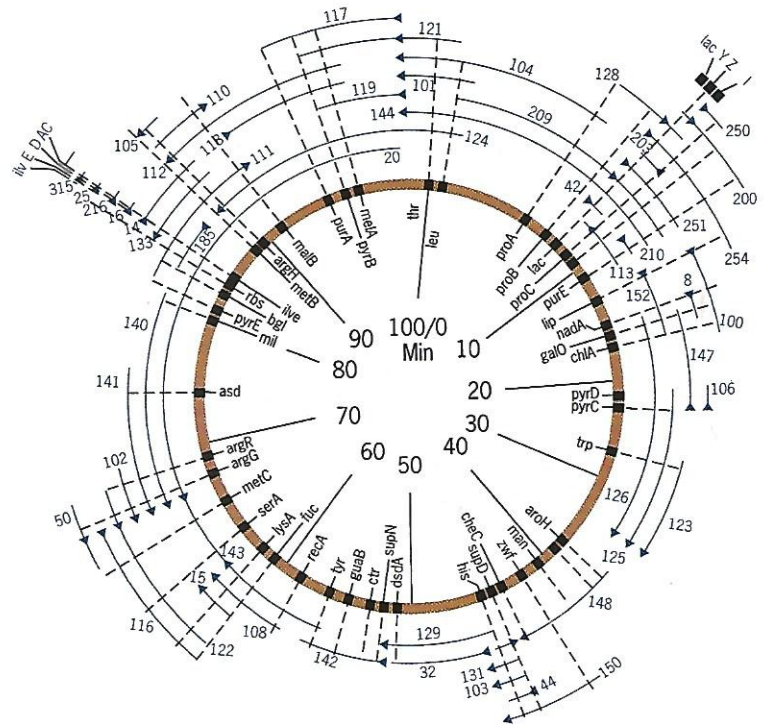
A transdução – outro mecanismo de transferência gênica em bactérias – foi descoberta por Norton Zinder e Joshua Lederberg em 1952. Zinder e Lederberg estudaram linhagens auxotróficas de *Salmonella typhimurium* cujo desenvolvimento exigia suplementos de aminoácidos. Uma linhagem necessitava de fenilalanina, triptofano e tirosina; a outra necessitava de metionina e histidina. Nenhuma linhagem era capaz de se desenvolver em meio mínimo sem esses aminoácidos. No entanto, quando Zinder e Lederberg cultivaram as linhagens juntas, foram produzidos raros prototróficos. Além disso, quando cultivaram as linhagens em meio contendo DNase, mas as separaram nos dois ramos do tubo em U (Figura 8.9), ainda foram produzidos recombinantes prototróficos. A insensibilidade à DNase exclui a transformação como mecanismo subjacente, e o fato de que o contato celular era dispensável para o surgimento dos prototróficos excluiu a conjugação. Experimentos subsequentes mostraram que uma das linhagens foi infectada por um vírus denominado bacteriófago P22 e que esse vírus levava genes de uma célula (doadora) para outra (receptora). Portanto, os raros prototróficos detectados por Zinder e Lederberg



■ **FIGURA 8.21** Formação de um *F'*. A excisão anômala do fator F de um cromossomo de célula Hfr produz um fator F que tem os genes *thr* e *leu* da *E. coli* e designado *F' thr leu*.

foram produzidos por recombinação entre o DNA bacteriano levado pelo vírus e o DNA no cromossomo da célula receptora.

Estudos posteriores mostraram que existem dois tipos muito diferentes de transdução. Na **transdução generalizada**, há na cabeça do fago um fragmento aleatório ou quase aleatório do DNA bacteriano em vez do cromossomo do fago. Na **transdução especializada**, há uma recombinação entre o cromossomo do hospedeiro e o cromossomo do fago, produzindo um cromossomo do fago que contém um trecho de DNA bacteriano. As



■ **FIGURA 8.22** *F'* em *E. coli*. Mapa do cromossomo de *E. coli* K12 mostrando os genes presentes em *F'* representativos. Os *F'* são desenhados como estruturas lineares para alinhá-los com os segmentos cromossômicos que contêm. Na realidade, são moléculas circulares de DNA – as estruturas formadas pela união das duas extremidades de cada *F'*.

Resolva!

Como mapear genes próximos usando diploides parciais?

Suponha que se queira determinar a ordem de dois genes (*y* e *z*) em um locus em relação a um marcador (*x*) em um locus próximo. Fazem-se os seguintes cruzamentos recíprocos:

1. célula doadora $x^+y^+z^- \times$ célula receptora $x^-y^-z^+$ e
2. célula doadora $x^-y^-z^+ \times$ célula receptora $x^+y^+z^-$.

Observe que a ordem dos três genes (*x*, *y* e *z*) é desconhecida; eles são escritos arbitrariamente em ordem alfabética. Suponha que todos os mutantes sejam auxotróficos e que se possam preparar meios seletivos nos quais cresçam apenas recombinantes prototróficos ($x^+y^+z^+$). Quando quantidades iguais de prole são plaqueadas em meio seletivo, observam-se cerca de 200 recombinantes prototróficos no cruzamento 1, enquanto se detectam mais de 4.000 no cruzamento 2. Qual é a ordem dos três genes no cromossomo?

► **Leia a resposta do problema no site**
<http://gen-io.grupogen.com.br>.

partículas do fago que contêm DNA bacteriano são denominadas *partículas transdutoras*. As partículas transdutoras generalizadas contêm apenas DNA bacteriano. As partículas transdutoras especializadas sempre contêm DNA do fago e da bactéria.

Transdução generalizada

Fagos transdutores generalizados são capazes de transportar qualquer gene bacteriano de uma célula para outra – daí o nome transdução generalizada. Os fagos transdutores generalizados mais conhecidos são P22 em *S. typhimurium* e P1 em *E. coli*. Apenas 1 a 2% das partículas de fago produzidas por bactérias infectadas por P22 ou P1 contêm DNA bacteriano, e apenas 1 a 2% do DNA transferido são incorporados ao cromossomo da célula receptora por recombinação. Assim, o processo é bastante ineficiente; a frequência de transdução de qualquer gene bacteriano é de aproximadamente 1 por 10^6 partículas de fago.

Transdução especializada

A transdução especializada é característica de vírus que transferem apenas determinados genes entre bactérias. O bacteriófago lambda (λ) é o fago transdutor

especializado mais conhecido; λ leva apenas os genes *gal* (necessário para uso da galactose como fonte de energia) e *bio* (essencial para a síntese de biotina) de uma célula de *E. coli* para outra. Já comentamos neste capítulo a inserção sítio-específica do cromossomo do λ no cromossomo da *E. coli* para criar um estado lisogênico (veja Bacteriófago lambda). O sítio de inserção está entre os genes *gal* e *bio* no cromossomo da *E. coli* (Figura 8.5), o que explica por que o λ transduz apenas esses genes.

O cromossomo de λ integrado – o prófago λ – em uma célula lisogênica sofre excisão espontânea raramente (cerca de uma em cada 10^5 divisões celulares), que causa sua entrada na via lítica. A excisão do prófago também pode ser induzida, por exemplo, por irradiação de células lisogênicas com luz ultravioleta. A excisão normal é basicamente o inverso do processo de integração sítio-específico e produz cromossomos circulares do fago e de bactérias intactos (Figura 8.23A). Às vezes, a

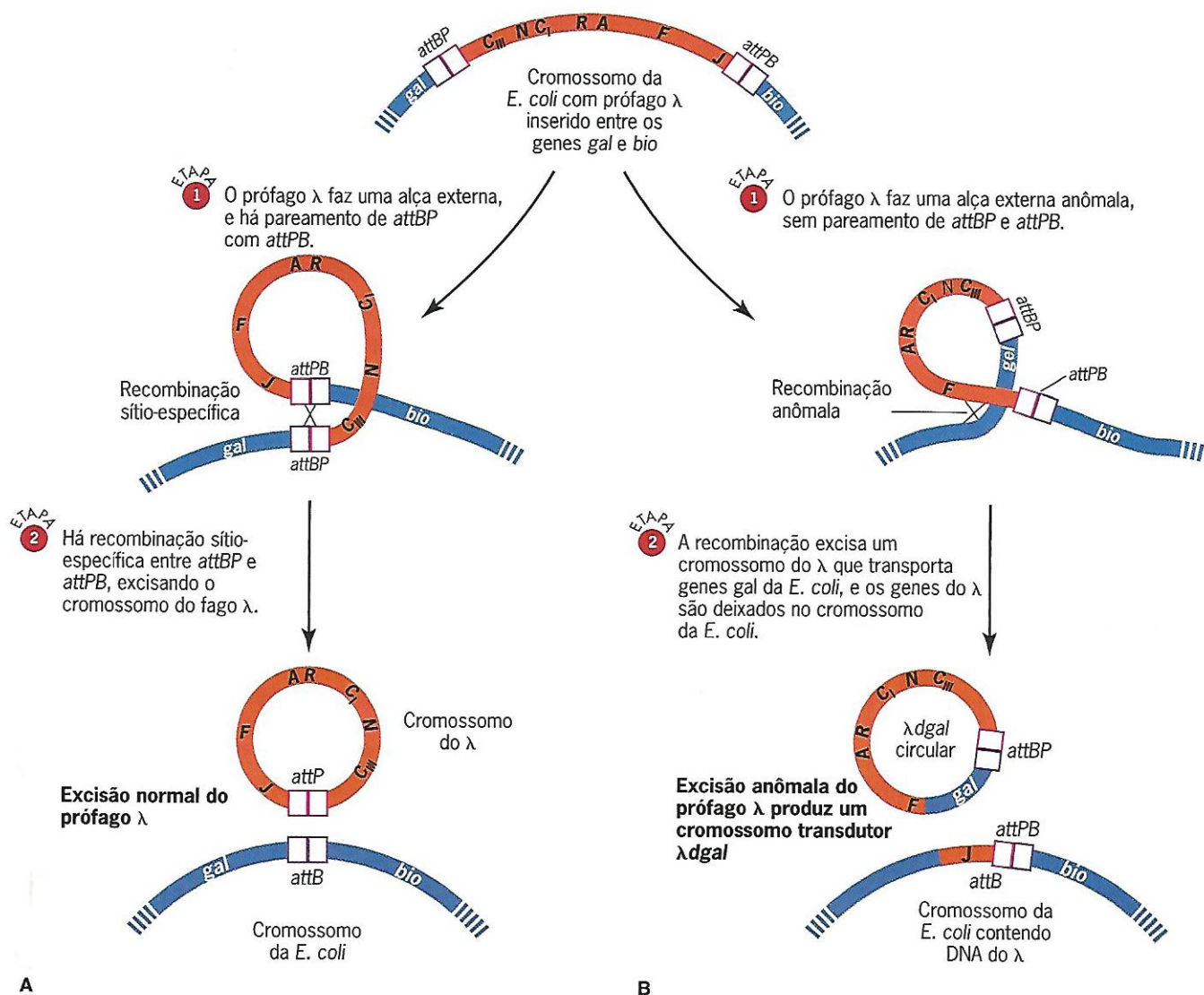


FIGURA 8.23 Excisão de prófago lambda. Comparação entre (A) excisão normal do prófago λ e (B) excisão anômala com produção de cromossomos transdutores λ gal.

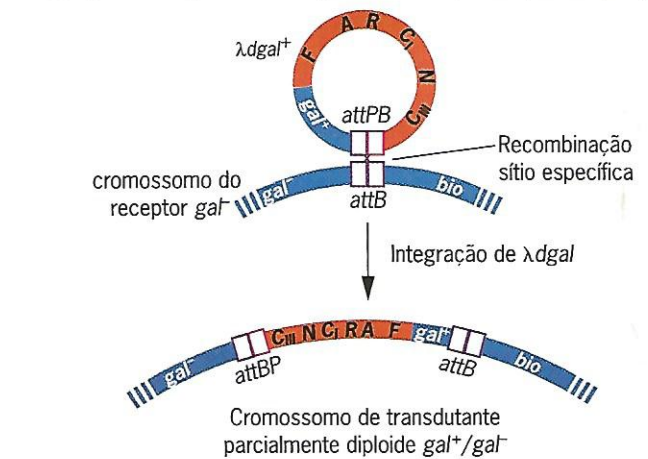
excisão é anômala, com ocorrência de *crossing over* em outro sítio que não o sítio de ligação original. Quando isso acontece, uma parte do cromossomo bacteriano é excisada com o DNA do fago e uma parte do cromossomo do fago é mantida no cromossomo do hospedeiro (Figura 8.23B). Essas excisões do prófago anômalo produzem fago transdutor especializado que leva os genes *gal* ou *bio* do hospedeiro. Esses fagos transdutores são designados λ *dg* (fago λ deficiente transportador de genes *gal*) e λ *dbio* (fago λ deficiente transportador de genes *bio*), respectivamente. Eles são partículas de fago deficientes porque um ou mais genes necessários para a reprodução lítica ou lisogênica permaneceram no cromossomo do hospedeiro.

Em vista do tamanho pequeno da cabeça do fago, apenas genes bacterianos situados próximos do prófago podem ser excisados com o DNA do fago e acondicionados na cabeça dos fagos. Outro fago transdutor especializado, 80, integra-se perto dos genes *trp* de *E. coli* (necessários para a síntese do aminoácido triptofano); esse fago transduz marcadores *trp*. Caso sejam formadas partículas transdutoras especializadas durante a excisão do prófago, como mostra a Figura 8.23B, elas só devem ser produzidas quando células lisogênicas entrarem na via lítica. Na verdade, as partículas transdutoras não estão presentes em lisados produzidos a partir de infecções líticas primárias. A frequência de partículas transdutoras em lisados produzidos por indução de células lisogênicas é de aproximadamente 1 em cada 10^6 partículas da prole; portanto, esses lisados são denominados *Lft* (do inglês, *low-frequency transduction*, transdução de baixa frequência).

O destino das moléculas de DNA λ *dg* e λ *bio* depois da injeção em novas células hospedeiras depende dos genes do λ ausentes. Se os genes para crescimento lítico estiverem ausentes, mas houver um sítio *att* e um gene *int* (integrase), os cromossomos defeituosos serão capazes de se integrar ao cromossomo do hospedeiro. No entanto, não serão capazes de se reproduzir liticamente, exceto se houver um λ de tipo selvagem, agindo como fago “auxiliar”. Se o gene *int* estiver ausente, o cromossomo do fago defeituoso só será capaz de se integrar na presença de um auxiliar de tipo selvagem. Se um fago λ *dg*⁺ infectar uma célula receptora *gal*⁻, a integração do λ *dg*⁺ produzirá um diploide parcial *gal*⁺/*gal*⁻ instável (Figura 8.24A), enquanto eventos de recombinação raros entre *gal*⁺ no DNA transdutor e *gal*⁻ no cromossomo receptor produzirão transdutantes *gal*⁺ (Figura 8.24B).

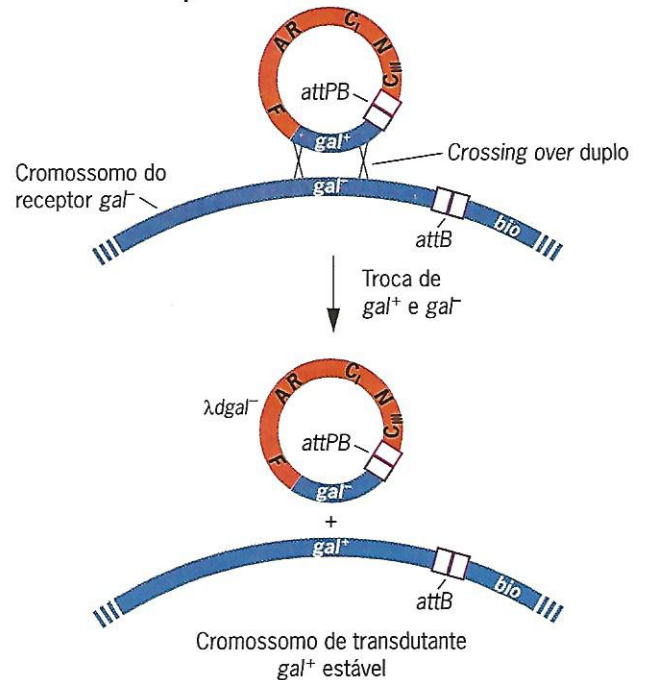
Se a proporção entre fagos e bactérias for alta, as células receptoras serão infectadas tanto pelo fago λ de tipo selvagem quanto por λ *dg*⁺; portanto, essas células serão lisogênicos duplos que transportam um prófago λ de tipo selvagem e um prófago λ *dg*. Os transdutantes produzidos serão diploides parciais *gal*⁺/*gal*⁻. Se os transdutantes *gal*⁺/*gal*⁻ forem induzidos com luz ultravioleta, os lisados conterão cerca de 50% de partículas λ *dg* e

Integração de λ *dg*⁺ em *attB* produz um diploide parcial *gal*⁺/*gal*⁻.



A

Um *crossing over* duplo insere o alelo *gal*⁺ de λ *dg*⁺ no cromossomo do hospedeiro.



B

FIGURA 8.24 Recombinação em células receptoras *gal*⁻ infectadas por fago transdutor λ *dg*⁺. A. Integração de λ *dg*⁺ em *attB* produz um diploide parcial *gal*⁺/*gal*⁻ instável. B. Um *crossing over* duplo transfere o alelo *gal*⁺ de λ *dg*⁺ para o cromossomo.

50% de partículas λ ⁺. Ambos os prófagos se replicarão com igual eficiência usando os produtos gênicos codificados pelo genoma λ ⁺. Esses lisados são denominados *Hft* (do inglês, *high-frequency transduction*, transdução de alta frequência). Os lisados *Hft* aumentam radicalmente a frequência dos eventos de transdução; portanto, lisados *Hft* são usados preferencialmente em experimentos de transdução.

PONTOS ESSENCIAIS

- Três processos parassexuados – transformação, conjugação e transdução – ocorrem em bactérias. Esses processos são distinguidos por dois critérios: a inibição da transferência gênica por desoxirribonuclease e a necessidade de contato celular
- A transformação implica a captação de DNA livre por bactérias
- A conjugação ocorre quando uma célula doadora faz contato com uma célula receptora e transfere DNA para a célula receptora
- A transdução ocorre quando um vírus leva genes bacterianos de uma célula doadora para uma célula receptora
- Os plasmídios são elementos genéticos extracromossômicos autorreplicantes
- Os episômos são replicados de maneira autônoma ou como componentes integrados de cromossomos bacterianos
- Os fatores F que contêm genes cromossômicos (fatores F') são transferidos para células F⁻ por sexodução.

Significado evolutivo da troca genética em bactérias

A troca genética é tão importante em bactérias quanto em outros organismos.

A mutação é a fonte de nova variação genética, e a recombinação produz novas combinações dessa variação – novas distribuições de genes responsáveis pelos fenótipos influenciados por seleção natural. Em eucariotos, os eventos de recombinação – distribuição independente e *crossing over* – são parte da reprodução sexuada. Mas a reprodução das bactérias não é sexuada. Apesar disso, a recombinação certamente é tão importante na evolução das bactérias quanto na evolução de eucariotos. Assim, não causa surpresa que mecanismos para troca de informações genéticas e recombinação de genes tenham se desenvolvido em bactérias. Esses mecanismos parassexuados – transformação, conjugação e transdução – geram novas combinações de genes e possibilitam a evolução e a adaptação das bactérias a novos nichos ambientais e a modificações súbitas dos *habitats* existentes. Teste seu conhecimento sobre a importância da recombinação em bactérias lendo Resolva!: Como evoluem os genomas bacterianos?

Embora os processos parassexuados sejam benéficos para bactérias, criam graves problemas para os seres humanos que estão tentando combater doenças bacterianas. Esses processos parassexuados são parcialmente responsáveis pela rápida evolução e disseminação de plasmídios que dotam as bactérias de resistência a antibióticos e fármacos. O uso disseminado de antibióticos na agricultura e na medicina resultou no surgimento de plasmídios que têm uma série completa de genes de resistência a antibióticos (veja O futuro: Bactérias resistentes a antibióticos).

Resolva!

Como evoluem os genomas bacterianos?

Um cruzamento (I) entre linhagens F⁺ *met*⁺ *ser*⁺ *cys*⁺ *str*^s e F⁻ *met*⁻ *ser*⁻ *cys*⁻ *str*^r de *E. coli* produziu todas as bactérias F⁺, mas não produziu recombinantes prototróficos *met*⁺ *ser*⁺ *cys*⁺ *str*^r. Depois de várias gerações, fizeram-se novas culturas de cada linhagem a partir de uma colônia, e o cruzamento foi repetido. Dessa vez (cruzamento II), foram produzidos recombinantes *met*⁺ *ser*⁺ *cys*⁺ *str*^r, mas todos esses recombinantes foram F⁻. Depois de várias outras gerações das linhagens usadas no cruzamento II, fizeram-se novas culturas a partir de colônias isoladas, e o cruzamento foi repetido pela terceira vez (cruzamento III). Não foram produzidos recombinantes *met*⁺ *ser*⁺ *cys*⁺ *str*^r no cruzamento III; em vez disso, toda a prole que sobreviveu no meio contendo estreptomicina tinha o genótipo *met*⁺ *ser*⁺ *cys*⁻ *str*^r e fenótipo F⁺. Usando um mapa do cromossomo da *E. coli*, explique esses resultados.

► Leia a resposta do problema no site <http://gen-io.grupogen.com.br>.

Na verdade, o ressurgimento da tuberculose na cidade de Nova York na década de 1990 foi, em grande parte, decorrente do surgimento de uma linhagem de *M. tuberculosis* resistente a sete antibióticos diferentes. Além disso, a linhagem de Nova York está intimamente relacionada com as linhagens predominantes de *M. tuberculosis* resistente na China e em outras regiões do mundo. Portanto, o que beneficia as bactérias patogênicas geralmente prejudica os seres humanos.

PONTOS ESSENCIAIS

- Os mecanismos parassexuados de recombinação produzem novas combinações de genes em bactérias
- Os mecanismos parassexuados estimulam a capacidade de adaptação das bactérias às alterações no ambiente.

O FUTURO

BACTÉRIAS RESISTENTES A ANTIBIÓTICOS

Em março de 2010, a Organização Mundial da Saúde divulgou que o aumento das linhagens multirresistentes (MDR, do inglês, *multiple drug-resistant*) de *Mycobacterium tuberculosis*, a bactéria causadora da tuberculose (TB), havia alcançado níveis recorde. Em alguns países, nossos melhores antibióticos são ineficazes no tratamento de até um quarto dos indivíduos com TB.

Agora, surgiu um novo gene, designado *NDM-1* (metalobeta-lactamase de Nova Delhi), que torna as bactérias resistentes a um importante grupo de antibióticos, os carbapenêns, usados com frequência no tratamento de linhagens MDR das bactérias. O gene *NDM-1* está localizado em um plasmídeo facilmente transferido de uma bactéria para outra. Até hoje, o gene foi encontrado em *E. coli* e *Klebsiella pneumoniae* – ambas causadoras de infecções urinárias graves – mas não restam muitas dúvidas de que vá se disseminar para outras espécies de bactérias. Infelizmente, estão sendo desenvolvidos apenas dois antibióticos com possibilidade de eficácia no tratamento das superbactérias *NDM-1*.

O que levou à evolução dessas bactérias resistentes a antibióticos e fármacos? Como os seres humanos contribuíram para essa possível crise? Podemos resolver esse problema? Em caso afirmativo, como? Primeiro, vamos analisar a história dos antibióticos e das bactérias resistentes a antibióticos.

Selman Waksman, ucraniano que imigrou para os EUA, descobriu a estreptomicina em 1943. Mais tarde, deu o nome de antibióticos a essa classe de fármacos antibacterianos. O primeiro tratamento documentado de um ser humano com estreptomicina foi o de uma mulher de 21 anos na Clínica Mayo, em Rochester, Minnesota. Essa paciente tinha um caso avançado de tuberculose. Em 1944 foi iniciado o tratamento experimental com injeções de estreptomicina e, para surpresa de todos, a tuberculose foi curada. Rapidamente, a estreptomicina e outros antibióticos tornaram-se “fármacos miraculosos”. A administração a pessoas com infecções bacterianas salvou milhões de vidas.

Logo os seres humanos começaram a usar grandes quantidades de antibióticos. Em 1950, o mundo usou 10 toneladas de estreptomicina. Em 1955, o uso mundial de estreptomicina havia aumentado para 50 t, mais cerca de 10 t de cloranfenicol e 10 t de tetraciclina.

No entanto, as bactérias logo começaram a reagir. Elas desenvolveram novos genes codificadores de produtos que as protegiam contra os antibióticos. O surgimento de bactérias resistentes a antibióticos confirmou o poder da seleção natural. Uma bactéria sem gene de resistência a antibióticos é destruída pelo antibiótico. Uma bactéria que tem o gene de resistência cresce, divide-se e produz uma população de bactérias, todas resistentes ao antibiótico. O resultado era inevitável: a disseminação de bactérias resistentes a antibióticos.

Alguns dos primeiros estudos que documentam a evolução de bactérias resistentes a antibióticos e fármacos foram realizados no Japão nas quatro “espécies” de *Shigella* – *S. dysenteriae*, *S. flexneri*, *S. boydii* e *S. sonnei* – causadoras de disenteria. Apenas 0,2% das

linhagens de *Shigella* isoladas em esgotos e rios poluídos em 1953 eram resistentes a algum dos antibióticos e fármacos testados. Apenas 12 anos mais tarde, a frequência de linhagens de *Shigella* resistentes a antibióticos e fármacos isoladas nos mesmos lugares aumentara para 58%. No entanto, a notícia verdadeiramente ruim não era a resistência dessas linhagens aos antibióticos, mas a resistência da maioria delas a no mínimo quatro dos seis antibióticos e fármacos – ampicilina, canamicina, tetraciclina, estreptomicina, sulfanilamida e cloranfenicol – testados. Eram linhagens multirresistentes de *Shigella*. Também começaram a surgir linhagens MDR de outras bactérias, como *M. tuberculosis*.

Os genes que protegiam as bactérias contra antibióticos geralmente estão presentes em pequenas moléculas de DNA denominadas plasmídios R (ver Plasmídios e episômos). Muitos plasmídios R são autotransmissíveis; isto é, têm genes que medeiam a própria transferência de uma célula para outra e até mesmo de uma espécie para outra. Além disso, os genes de resistência a antibióticos costumam estar presentes em elementos genéticos – elementos transponíveis ou “elementos genéticos móveis” – que podem se mover de uma molécula de DNA para outra (Capítulo 17). Assim, os genes nesses plasmídios R podem se disseminar rapidamente nas populações de bactérias.

Uma explicação para o desenvolvimento tão rápido de linhagens MDR de bactérias é o uso excessivo de antibióticos. Muitas vezes, prescrevem-se antibióticos contra infecções virais como o resfriado comum e a gripe. Os antibióticos não têm atividade antiviral e não devem ser usados no tratamento das infecções virais. Além disso, os antibióticos são usados em larga escala e grandes quantidades como “promotores do crescimento” na alimentação de animais para evitar infecções bacterianas que reduzem a velocidade de crescimento. Na verdade, quase metade dos antibióticos produzidos nos EUA é usada como aditivo na ração de animais. Eles são acrescentados em proporções de 2 a 50 g por tonelada de ração, e ocorre o inevitável – surgem bactérias resistentes a antibióticos. Então, essas bactérias resistentes são transmitidas aos seres humanos que cuidam dos animais, trabalham na indústria de embalagem de carnes ou consomem carne mal passada.

Em face do surgimento disseminado de linhagens MDR de *M. tuberculosis*, *Staphylococcus aureus*, *Shigella dysenteriae* e outras bactérias patogênicas, talvez seja recomendável restringir o uso de alguns de nossos melhores antibióticos ao tratamento de doenças humanas possivelmente fatais. Na verdade, a Dinamarca proibiu o uso de penicilinas e tetraciclina como promotores do crescimento de animais na década de 1970, e a Suécia proibiu qualquer uso não terapêutico de antibióticos, inclusive como promotores do crescimento de animais, em 1986. Os efeitos negativos da proibição do uso de antibióticos na ração de animais sobre a produtividade foi mínimo, e, na Suécia, o uso geral de antibióticos caiu 55% desde o início da proibição de usos não terapêuticos. Talvez seja hora de os EUA e o restante do mundo seguirem o exemplo da Escandinávia – proibir, ou ao menos limitar, o uso não terapêutico de antibióticos. Realmente precisamos acrescentar antibióticos à ração dos animais? E no sabonete, precisamos dele?

Exercícios

Aplique a análise genética básica

- Quais as vantagens dos vírus em relação a organismos celulares e multicelulares para a pesquisa genética?

Resposta: As duas principais vantagens dos vírus em relação a organismos celulares e multicelulares para os estudos genéticos são (1) a simplicidade estrutural e (2) o ciclo de vida curto. Em geral, os vírus têm um único cromossomo, com um número relativamente pequeno de genes, e podem completar o ciclo de vida em um período que varia de cerca de 20 min a algumas horas.

- Quais são as principais diferenças entre o *crossing over* em bactérias e eucariotos?

Resposta: O *crossing over* em bactérias geralmente ocorre entre um fragmento do cromossomo de uma célula doadora e um cromossomo circular intacto em uma célula receptora (Figura 8.7A). Desse modo, os *crossing overs* têm de ocorrer em números pares, que inserem segmentos do cromossomo da célula doadora no cromossomo da célula receptora. O *crossing over* simples, ou qualquer número ímpar de *crossing overs*, destrói a integridade do cromossomo circular e deixa em seu lugar uma molécula linear de DNA (Figura 8.7B).

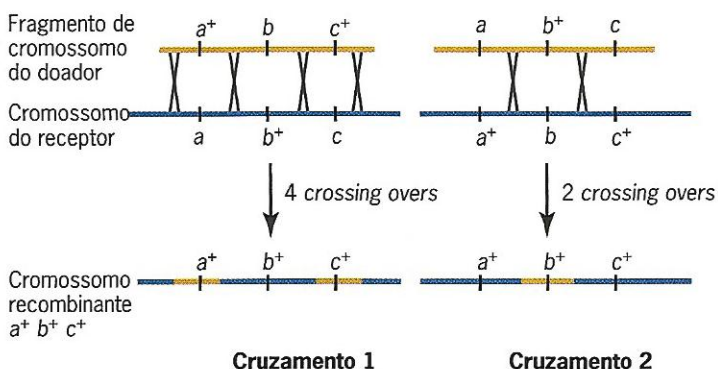
- Quando cultivadas juntas, duas linhagens de *E. coli*, $a^+ b^+$ e $a^+ b^-$, trocam material genético, levando à produção de recombinantes $a^+ b^+$. No entanto, quando essas duas linhagens são cultivadas em ramos opostos de um tubo em U (Figura 8.9), não há produção de recombinantes $a^+ b^+$. Que processo parassexual é responsável pela formação dos recombinantes $a^+ b^+$ quando essas linhagens são cultivadas juntas?

Resposta: As duas linhagens de *E. coli* estão trocando informações por conjugação, o único processo parassexual em bactérias que requer contato celular.

O filtro de vidro que separa os ramos do tubo em U impede o contato entre células nesses ramos.

- Você identificou três marcadores genéticos próximos – a , b e c – em *E. coli*. Os marcadores são transferidos de uma linhagem Hfr para uma linhagem F^- em menos de 1 min, e estão presentes no cromossomo na ordem $a-b-c$. Você faz experimentos de transdução com o fago P1 usando linhagens de genótipo $a^+ b^+ c^+$ e $a^+ b^- c^+$. No cruzamento 1, as células doadoras são $a^+ b^+ c^+$ e as células receptoras são $a^+ b^- c^+$. No cruzamento 2, as células doadoras são $a^+ b^- c^+$ e as células receptoras são $a^+ b^+ c^+$. Nos dois cruzamentos, você prepara placas com meio mínimo em que apenas os recombinantes $a^+ b^+ c^+$ produzem colônias. Em que cruzamento você esperaria observar a maioria dos recombinantes $a^+ b^+ c^+$?

Resposta: Você esperaria mais recombinantes $a^+ b^+ c^+$ no cruzamento 2 porque a formação de um cromossomo com os três marcadores de tipo selvagem só requer dois *crossing overs* (um par de *crossing overs*) nesse cruzamento, ao passo que são necessários quatro *crossing overs* (dois pares) para produzir um cromossomo $a^+ b^+ c^+$ no cruzamento 1. Os *crossing overs* necessários são mostrados no diagrama a seguir.



Autoavaliação

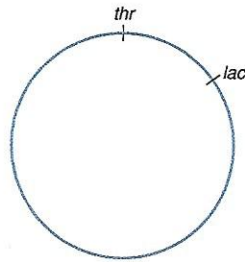
Integre diferentes conceitos e técnicas

- Você identificou uma linhagem mutante de *E. coli* que não sintetiza histidina (His^-). Com o objetivo de determinar a localização da mutação his^- no cromossomo de *E. coli*, você faz experimentos de cruzamento interrompido com cinco linhagens diferentes de Hfr. O diagrama a seguir mostra o momento de entrada (minutos, entre parênteses) dos alelos selvagens dos cinco primeiros marcadores (genes mutantes) na linhagem His^- .

Hfr A —	<i>bio</i> (4)	<i>glu</i> (20)	<i>his</i> (27)	<i>cys</i> (37)	<i>tyr</i> (45)
Hfr B —	<i>xyl</i> (6)	<i>met</i> (18)	<i>tyr</i> (24)	<i>cys</i> (32)	<i>his</i> (42)
Hfr C —	<i>his</i> (3)	<i>cys</i> (13)	<i>tyr</i> (21)	<i>met</i> (27)	<i>xyl</i> (39)
Hfr D —	<i>xyl</i> (7)	<i>thr</i> (25)	<i>lac</i> (40)	<i>bio</i> (48)	<i>glu</i> (62)
Hfr E —	<i>his</i> (4)	<i>glu</i> (11)	<i>bio</i> (27)	<i>lac</i> (35)	<i>thr</i> (50)

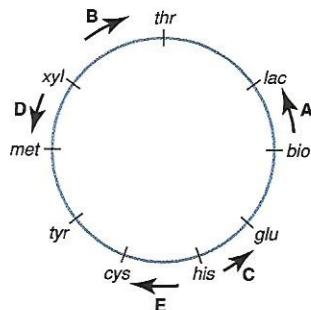
- No mapa a seguir do cromossomo circular da *E. coli*, indique (1) a localização relativa de cada

gene, (2) a posição em que o fator F é integrado a cada um dos cinco Hfr e (3) a direção da transferência de cromossomo para cada Hfr (indique a direção com uma seta).



(b) Para definir melhor a localização da mutação *his*⁻ no cromossomo, você usa a linhagem mutante como receptora em um experimento de transdução com bacteriófago P1. Considerando-se que o fago P1 comporta cerca de 1% da molécula de DNA cromossômico de *E. coli*, você esperaria que houvesse cotransdução de algum dos genes mostrados na figura anterior com o alelo *his*⁺ de seu gene mutante *his*? Em caso afirmativo, quais deles? Observe que o cromossomo de *E. coli* contém 4,6 milhões de pares de nucleotídeos e que a transferência de todo o cromossomo durante a conjugação leva 100 min. Justifique sua resposta.

Resposta: (a) A ordem dos genes é mostrada no mapa adiante, e os sítios de integração do fator F e a direção de transferência em cada Hfr são indicados pelas setas A a E.



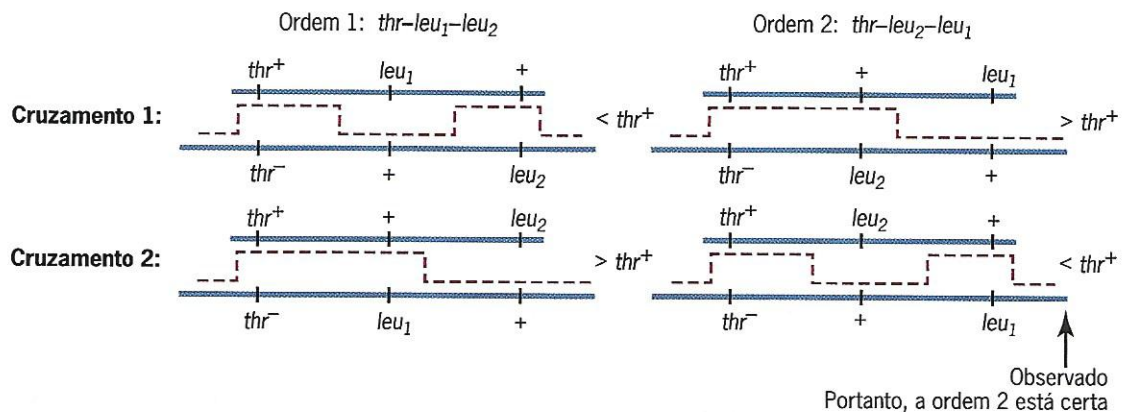
(b) Não haveria cotransdução de nenhum marcador com *his*⁺ porque o fago P1 só comporta 1% do cromossomo de *E. coli*, e nenhum dos outros genes está dentro de 1 min de *his*.

2. Cruzamentos com transdução de três pontos recíprocos foram usados para determinar a ordem de duas mutações, *leu*₁ e *leu*₂, no gene *leuA* em relação ao gene ligado *thrA* da *E. coli*. Em cada cruzamento, recombinantes *leu*⁺ foram selecionados em meio mínimo contendo treonina, mas não leucina, e testados para *thr*⁺ ou *thr*⁻ por plaqueamento em réplica em placas que não continham treonina. Os resultados são apresentados na tabela adiante:

Cruzamento			
Marcadores do doador	Marcadores do receptor	Alelo <i>thr</i> em recombinantes <i>leu</i> ⁺	Porcentagem de <i>thr</i> ⁺
1. <i>thr</i> ⁺ <i>leu</i> ₁	<i>thr</i> ⁻ <i>leu</i> ₂	350 <i>thr</i> ⁺ : 349 <i>thr</i> ⁻	50
2. <i>thr</i> ⁺ <i>leu</i> ₂	<i>thr</i> ⁻ <i>leu</i> ₁	60 <i>thr</i> ⁺ : 300 <i>thr</i> ⁻	17

Qual é a ordem de *leu*₁ e *leu*₂ em relação ao marcador externo *thr*?

Resposta: Os diagramas dos dois cruzamentos são apresentados, mostrando as duas ordens possíveis, e as linhas vermelhas tracejadas marcam as partes dos dois cromossomos que têm de estar presentes em recombinantes *thr*⁺-*leu*₁⁺-*leu*₂⁺ (+ + +). Note que se a ordem 1 estiver certa, a formação de recombinantes + + + exigirá 4 *crossing overs* (2 pares de *crossing overs*) no cruzamento 1 e apenas 2 *crossing overs* (1 par) no cruzamento 2, portanto, prevenindo mais recombinantes + + + no cruzamento 2 e menos no cruzamento 1. Mas se a ordem 2 estiver certa, deve haver mais recombinantes + + + no cruzamento 1 e menos no cruzamento 2. Visto que foi observado o segundo resultado, a ordem certa é *thr*-*leu*₂-*leu*₁.



Avaliação adicional

Entenda melhor e desenvolva a capacidade analítica

- 8.1 Quais critérios consideram os vírus seres vivos? Não vivos?
- 8.2 Qual é a diferença entre bacteriófagos e outros vírus?
- 8.3 Quais são as diferenças entre o ciclo de vida dos bacteriófagos T4 e λ ? Quais são as semelhanças?
- 8.4 Quais são as diferenças entre as estruturas do pró-fago λ e do cromossomo do λ acondicionado no vírion λ ?
- 8.5 Qual é a diferença entre a integração do cromossomo do λ ao cromossomo do hospedeiro durante uma infecção lisogênica e o *crossing over* entre cromossomos homólogos?
- 8.6 Os geneticistas usaram mutações que causam fenótipos alterados como olhos brancos em *Drosophila*, flores brancas e sementes rugosas em ervilhas, e alteração da cor da pelagem em coelhos para determinar as localizações dos genes nos cromossomos desses eucariotos. Que tipos de fenótipos mutantes foram usados para mapear genes em bactérias?
- 8.7 Você identificou três mutações – *a*, *b* e *c* – em *Streptococcus pneumoniae*. Todos os três são recessivos para seus alelos selvagens a^+ , b^+ e c^+ . Você prepara o DNA a partir de uma linhagem doadora de tipo selvagem e a usa para transformar uma linhagem com genótipo *a b c*. Você observa transformantes a^+ b^+ e transformantes a^+ c^+ , mas não transformantes b^+ c^+ . Essas mutações estão próximas? Em caso afirmativo, qual é sua ordem no cromossomo do *Streptococcus*?
- 8.8 Uma linhagem de *E. coli* com deficiência nutricional cresce apenas em meio contendo timina, enquanto outra linhagem com deficiência nutricional cresce apenas em meio contendo leucina. Quando essas duas linhagens são cultivadas juntas, alguns organismos da prole são capazes de crescer em meio mínimo que não contém timina nem leucina. Como é possível explicar esse resultado?
- 8.9 Suponha que você acabou de demonstrar a recombinação genética (p. ex., quando uma linhagem de genótipo *a b^+* está presente em uma linhagem de genótipo a^+ *b*, formam-se alguns genótipos recombinantes, a^+ b^+ e *a b*) em uma espécie de bactérias não estudada antes. Como você determinaria se a recombinação observada foi resultado de transformação, conjugação ou transdução?
- 8.10 (a) Quais são as diferenças genotípicas entre células F^- , células F^+ e células Hfr? (b) Quais são as diferenças fenotípicas? (c) Qual é o mecanismo de conversão das células F^- em células F^+ ? Das células F^+ em Hfr? Das células Hfr em F^+ ?
- 8.11 (a) Qual é a utilidade dos fatores F' em análise genética? (b) Como são formados os fatores F' ? (c) Qual é o mecanismo de sexodução?
- 8.12 Quais são as diferenças básicas entre transdução generalizada e transdução especializada?
- 8.13 Que papéis representam os elementos IS na integração de fatores F ?
- 8.14 Como é possível mapear os genes bacterianos por experimentos de cruzamento interrompido?
- 8.15 O que é cotransdução? Como se podem usar as frequências de cotransdução para mapear marcadores genéticos?
- 8.16 Em *E. coli*, a capacidade de usar lactose como fonte de carbono requer a presença das enzimas β -galactosidase e β -galactosídeo permease. Essas enzimas são codificadas por dois genes próximos, *lacZ* e *lacY*, respectivamente. Outro gene, *proC*, controla, em parte, a capacidade das células de *E. coli* de sintetizar o aminoácido prolina. Os alelos *str^r* e *str^s*, respectivamente, controlam a resistência e a sensibilidade à estreptomicina. Sabe-se que Hfr H transfere os dois genes *lac*, *proC* e *str*, nessa ordem, durante a conjugação.
- Fez-se um cruzamento entre Hfr H de genótipo *lacZ^- lacY^+ proC^+ str^s* e uma linhagem F^- de genótipo *lacZ^+ lacY^- proC^- str^r*. Depois de aproximadamente 2 h, a mistura foi diluída e plaqueada em meio que continha estreptomicina, mas não prolina. Quando se analisou a capacidade das colônias recombinantes *proC^+ str^r* de crescer em meio contendo lactose como única fonte de carbono, pouquíssimas delas foram capazes de fermentar a lactose. Quando se fez o cruzamento recíproco (Hfr H *lacZ^- lacY^+ proC^+ str^s* \times F^- *lacZ^+ lacY^- proC^- str^r*), muitos dos recombinantes *proC^+ str^r* foram capazes de crescer em meio contendo lactose como única fonte de carbono. Qual é a ordem dos genes *lacZ* e *lacY* em relação a *proC*?
- 8.17 Uma linhagem F^+ , marcada em 10 loci, dá origem espontaneamente à prole Hfr sempre que o fator F é incorporado ao cromossomo da linhagem F^+ . O fator F pode se integrar ao cromossomo circular em muitos pontos, de maneira que as linhagens Hfr resultantes transfiram os marcadores genéticos em ordens diferentes. Em qualquer linhagem Hfr, a ordem de marcadores que entram em uma célula receptora pode ser determinada por experimentos de cruzamento interrompido. A partir dos dados

para várias linhagens Hfr derivadas da mesma linhagem F^+ , determine a ordem dos marcadores na linhagem F^+ .

Linhagem Hfr	Marcadores doados em ordem
1	– Z-H-E-R →
2	– O-K-S-R →
3	– K-O-W-I →
4	– Z-T-I-W →
5	– H-Z-T-I →

8.18 Os dados da tabela adiante foram obtidos a partir de testes de transdução de três pontos feitos para determinar a ordem de sítios mutantes no gene *A* que codifica a subunidade α do triptofano sintetase em *E. coli*. *Anth* é um marcador ligado não selecionado. Em cada cruzamento, recombinantes trp^+ foram selecionados e, depois, classificados em relação ao marcador *anth* ($anth^+$ ou $anth^-$). Qual é a ordem linear dos alelos *anth* e dos três alelos mutantes do gene *A* indicada pelos dados na tabela?

Cruzamento	Marcadores do doador	Marcadores do receptor	Alelo <i>anth</i> em recombinantes trp^+	% $anth^+$
1	$anth^+$ – A34	$anth^-$ – A223	72 $anth^+$: 332 $anth^-$	18
2	$anth^+$ – A46	$anth^-$ – A223	196 $anth^+$: 180 $anth^-$	52
3	$anth^+$ – A223	$anth^-$ – A34	380 $anth^+$: 379 $anth^-$	50
4	$anth^+$ – A223	$anth^-$ – A46	60 $anth^+$: 280 $anth^-$	20

8.19 O bacteriófago P1 medeia a transdução generalizada em *E. coli*. Um lisado transdutor de P1 foi preparado pelo cultivo do fago P1 em bactérias $pur^+ pro^- his^-$. Os genes *pur*, *pro* e *his* codificam enzimas necessárias para a síntese de purinas, prolina e histidina, respectivamente. Permitiu-se que o fago e as partículas transdutoras nesse lisado infectassem células $pur^- pro^+ his^+$. Depois de incubar as bactérias infectadas por um período suficiente para permitir que haja transdução, elas foram plaqueadas em meio mínimo suplementado com prolina e histidina, mas sem purinas para selecionar transdutantes pur^+ . As colônias pur^+ foram transferidas para meio mínimo com e sem prolina e com e sem histidina para determinar as frequências de cada marcador externo. Em face dos resultados a seguir, qual é a ordem dos três genes no cromossomo de *E. coli*?

Genótipo	Número observado
$pro^+ his^+$	100
$pro^- his^+$	22
$pro^+ his^-$	150
$pro^- his^-$	1

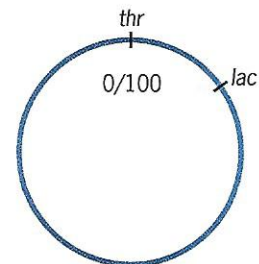
8.20 Duas outras mutações o gene *trp A* de *E. coli*, *trp A58* e *trp A487*, foram ordenadas em relação a *trp A223* e ao marcador externo *anth* por cruzamentos com transdução de três fatores, conforme descrito no Problema 8.18. Os resultados desses cruzamentos são resumidos na tabela a seguir. Qual é a ordem linear de *anth* e dos três sítios mutantes no gene *trp A*?

Cruzamento	Marcadores do doador	Marcadores do receptor	Alelo <i>anth</i> em recombinantes trp^+	% $anth^+$
1	$anth^+$ – A487	$anth^-$ – A223	72 $anth^+$: 332 $anth^-$	82
2	$anth^+$ – A58	$anth^-$ – A223	196 $anth^+$: 180 $anth^-$	48
3	$anth^+$ – A223	$anth^-$ – A487	380 $anth^+$: 379 $anth^-$	50
4	$anth^+$ – A223	$anth^-$ – A58	60 $anth^+$: 280 $anth^-$	80

8.21 Você identificou uma linhagem mutante de *E. coli* que não sintetiza histidina (His^-). Com o objetivo de determinar a localização da mutação *his^-* no cromossomo de *E. coli*, você faz experimentos de cruzamento interrompido com cinco linhagens diferentes de Hfr. O quadro a seguir mostra o momento de entrada (minutos, entre parênteses) dos alelos selvagens dos cinco primeiros marcadores (genes mutantes) na linhagem *His*.

Hfr A —	<i>his</i> (1)	<i>man</i> (9)	<i>gal</i> (28)	<i>lac</i> (37)	<i>thr</i> (45)
Hfr B —	<i>man</i> (15)	<i>his</i> (23)	<i>cys</i> (38)	<i>ser</i> (42)	<i>arg</i> (49)
Hfr C —	<i>thr</i> (3)	<i>lac</i> (11)	<i>gal</i> (20)	<i>man</i> (39)	<i>his</i> (47)
Hfr D —	<i>cys</i> (3)	<i>his</i> (18)	<i>man</i> (26)	<i>gal</i> (45)	<i>lac</i> (54)
Hfr E —	<i>thr</i> (4)	<i>rha</i> (18)	<i>arg</i> (36)	<i>ser</i> (43)	<i>cys</i> (47)

No mapa a seguir do cromossomo circular de *E. coli*, indique (1) a localização de cada gene em relação a *thr* (localizado em 0/100 min), (2) a posição em que o fator F é integrado a cada uma das cinco células Hfr e (3) a direção da transferência de cromossomo para cada Hfr (indique a direção com uma seta).



8.22 Sabe-se que as mutações *nrd 11* (gene *nrd B*, que codifica a subunidade beta da enzima ribonucleotídeo redutase), *am M69* (gene 63, que codifica uma proteína que auxilia a fixação da fibra da cauda) e *nd 28* (*denA*, que codifica a enzima endonuclease II) estão localizadas entre os genes 31 e 32 no

cromossomo do bacteriófago T4. As mutações *am* N54 e *am* A453 estão localizadas nos genes 31 e 32, respectivamente. Em vista dos dados do cruzamen-

to de três fatores na tabela a seguir, qual é a ordem linear dos cinco sítios mutantes?

Dados do cruzamento de três fatores

Cruzamento	% de recombinação ^a
1. <i>am</i> A453– <i>am</i> M69 × <i>nrd</i> 11	2,6
2. <i>am</i> A453– <i>nrd</i> 11 × <i>am</i> M69	4,2
3. <i>am</i> A453– <i>am</i> M69 × <i>nd</i> 28	2,5
4. <i>am</i> A453– <i>nd</i> 28 × <i>am</i> M69	3,5
5. <i>am</i> A453– <i>nrd</i> 11 × <i>nd</i> 28	2,9
6. <i>am</i> A453– <i>nd</i> 28 × <i>nrd</i> 11	2,1
7. <i>am</i> N54– <i>am</i> M69 × <i>nrd</i> 11	3,5
8. <i>am</i> N54– <i>nrd</i> 11 × <i>am</i> M69	1,9
9. <i>am</i> N54– <i>nd</i> 28 × <i>am</i> M69	1,7
10. <i>am</i> N54– <i>am</i> M69 × <i>nd</i> 28	2,7
11. <i>am</i> N54– <i>nd</i> 28 × <i>nrd</i> 11	2,9
12. <i>am</i> N54– <i>nrd</i> 11 × <i>nd</i> 28	1,9

^aTodas as frequências de recombinação são calculadas como $\frac{2 \text{ (prole de tipo selvagem)}}{\text{prole total}} \times 100$.

Genômica na Web em <http://www.ncbi.nlm.nih.gov>

O genoma de *E. coli* foi um dos primeiros genomas bacterianos sequenciados. A sequência nucleotídica completa (4,6 milhões de pares de nucleotídeos) do genoma da linhagem K12 de *E. coli* foi publicada em setembro de 1997.

1. Quantas linhagens diferentes de *E. coli* tiveram os genomas sequenciados desde 1997?
2. Todos esses genomas têm o mesmo tamanho aproximado? Caso não tenham, qual é a variação de tamanho observada entre os genomas de diferentes linhagens de *E. coli*?

3. Algumas linhagens de *E. coli*, por exemplo, 0157:H7, são mais patogênicas para seres humanos e outros mamíferos que linhagens como a K12. O genoma dessas linhagens é maior ou menor que o de K12? As comparações dos genes nas linhagens patogênicas e não patogênicas poderiam dar pistas sobre a razão de algumas linhagens serem patogênicas e outras não?

Dica: No site do NCBI, *Genome Biology* → *Entrez Genome* → *Microbial Genomes* → *Complete Genomes* → *Escherichia coli*.

ordem consistente com os dados observados é centrômero – *A* – *B*. Em seguida precisamos determinar que tipo de *crossing* duplo ocorreu no evento meiótico que produziu o asco padrão 4. Os diagramas dos *crossings* duplos possíveis mostram que o padrão tipo 4 resulta de um *crossing* duplo bifilamentar, com um *crossing* no intervalo I e um no intervalo II.

Para calcular as distâncias entre o centrômero e o gene *A* e entre os genes *A* e *B*, lembre-se de que nós já sabemos (1) que o asco padrão 2 resulta de um único *crossing* no intervalo II, (2) que o asco padrão 3 resulta de um único *crossing* no intervalo I e (3) que o asco padrão 4 resulta de um *crossing* duplo, com um *crossing* em cada intervalo. A distância de mapa do locus *A* para o locus *B* será de $[1/2 (\text{número de ascas do tipo 2} + \text{número de ascas do tipo 4}) / (\text{número total de ascas})] \times 100$, ou

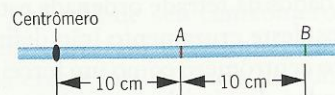
$$\frac{1/2(48 + 2)}{200} \times 100 = 12,5 \text{ unidades de mapa}$$

Os 50 ascas envolvidos neste cálculo resultaram todos de eventos meióticos nos quais houve um *crossing* no intervalo entre os loci *A* e *B*. Assim, 12,5 por cento da prole tem combinações recombinantes dos alelos dos genes *A* e *B*.

Similarmente, a distância do locus *A* até o centrômero será

$$\frac{1/2(38 + 2)}{200} \times 100 = 10 \text{ unidades de mapa}$$

Novamente, os 40 ascas contados neste cálculo resultaram todos de eventos meióticos nos quais havia um *crossing* no intervalo entre o locus *A* e o centrômero. A multiplicação pela metade corrige o fato de que cada *crossing* envolve apenas duas das quatro cromátides em uma tétrade. Apenas no caso dos ascas ditipo não parentais (resultantes de *crossings* duplos de quatro filamentos), o fator metade não é usado. Assim, o mapa de ligação definido por nossos dados hipotéticos de tétrade é



PONTOS IMPORTANTES

- Os genótipos de todos os quatro produtos de meiose, a tétrade, podem ser analisados em ascomicetos tais como a levedura *Saccharomyces cerevisiae* e o mofo do pão *Neurospora crassa*.
- As tétrades desordenadas de leveduras podem ser usadas para detectar ligação e mapear genes.
- As tétrades ordenadas de *Neurospora crassa* podem ser usadas para mapear as posições tanto dos centrômeros quanto dos genes em cromossomos.

Exercícios Básicos Ilustram a análise genética básica.

1. Que vantagens para a pesquisa genética têm os vírus em relação a organismos celulares e multicelulares?

Resposta: As duas principais vantagens para estudos genéticos que os vírus têm em relação a organismos celulares e multicelulares são (1) sua simplicidade estrutural e (2) seu curto ciclo de vida. Os vírus geralmente contêm um único cromossomo com um número relativamente pequeno de genes e podem completar seu ciclo de vida em cerca de 20 minutos a algumas horas. Cruzamentos do fago T4 podem ser feitos em menos de uma hora, e os resultados podem ser analisados no mesmo dia.

2. Como um cruzamento é feito com bacteriófagos?

Resposta: Um cruzamento de fagos é feito infectando-se simultaneamente as mesmas células hospedeiras com dois fagos de interesse (em geral mutantes) e triando-se quanto a recombinantes (geralmente tipo selvagem) entre a prole. A proporção de fagos para bactérias é mantida bem alta (cerca de 5 a 10 fagos de cada genótipo para cada bactéria) de modo que a maioria das células são infectadas com ambos os fagos parentais. Uma grande vantagem do bacteriófago em tais estudos genéticos é que um grande número de prole (10^8 a 10^9) pode facilmente ser triado quanto a recombinantes raros.

3. Quais são as principais diferenças entre *crossing-over* em bactérias e em eucariontes?

Resposta: Em bactérias, o *crossing-over* geralmente ocorre entre um fragmento do cromossomo de uma célula doadora e um cromossomo circular intacto em uma célula receptora (veja Fig. 8.15a). Como resultado, os *crossings* devem ocorrer aos pares que inserem segmentos do cromossomo da célula doadora no cromossomo da célula receptora. *Crossings* únicos, ou qualquer número ímpar de *crossings*, irão destruir a integridade do cromossomo circular e produzir uma molécula de DNA linear em seu lugar (veja Fig. 8.15b).

4. Quando cultivadas juntas, sabe-se que duas linhagens de *E. coli*, $a^+ b^+$ e $a^- b^-$, trocam material genético, resultando na produção de recombinantes $a^+ b^+$. Entretanto, quando estas duas linhagens são cultivadas em braços opostos de um tubo em U (veja Fig. 8.17), nenhum recombinante $a^+ b^+$ é produzido. Que processo parassexual é responsável pela formação de recombinantes $a^+ b^+$ quando tais linhagens são cultivadas juntas?

Resposta: As duas linhagens de *E. coli* estão trocando informações por conjugação, o único processo parassexual em bactérias que requer contato celular. O filtro de vidro que separa os braços do tubo em U impede contato entre células nestes braços.

5. Um geneticista de leveduras analisou 100 tétrades em um cruzamento envolvendo dois genes; 60 das tétrades são ditipos parentais, 8 são ditipos não parentais e 32 são tetratipos. (a) Que fato

estabelece que os dois genes estão ligados? (b) Qual a distância de mapa entre estes genes?

Resposta: (a) Para genes ligados, as frequências dos ascos ditipo parental e ditipo não parental devem ser aproximadamente iguais. Sua desigualdade óbvia nestes dados indica que os genes estão ligados. (b) A frequência de recombinação entre os genes é estimada pela fórmula $[(1/2)T + \text{NPD}]/\text{total}$, onde T

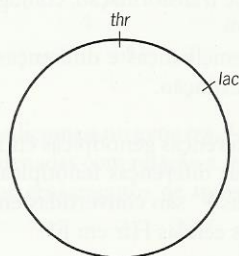
e NPD são o número de ascos tetratipo e ditipo não parental observados. A frequência de recombinação é portanto 0,24. Podemos usar este número para estimar a distância genética em centiMorgans, isto é, 24 centiMorgans, ou podemos ajustar o número para cima como objetivo de considerar múltiplas trocas pela fórmula $[(1/2)T + 3 \text{ NPD}]/\text{total}$, que produz uma distância de mapa de 40 centiMorgans.

Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. Você identificou um mutante de linhagem de *E. coli* que não pode sintetizar histidina (His^-). Para determinar a localização da mutação *bis*⁻ no cromossomo de *E. coli*, você faz experimentos de cruzamento interrompido com cinco linhagens Hfr diferentes. O mapa seguinte mostra o tempo de entrada (minutos, entre parênteses) dos alelos tipo selvagem dos primeiros cinco marcadores (genes mutantes) na linhagem His^- .

Hfr A — <i>bio</i> (4)	<i>glu</i> (20)	<i>bis</i> (27)	<i>cys</i> (37)	<i>tyr</i> (45)
Hfr B — <i>xyl</i> (6)	<i>met</i> (18)	<i>tyr</i> (24)	<i>cys</i> (32)	<i>bis</i> (42)
Hfr C — <i>bis</i> (3)	<i>cys</i> (13)	<i>tyr</i> (21)	<i>met</i> (27)	<i>xyl</i> (39)
Hfr D — <i>xyl</i> (7)	<i>thr</i> (25)	<i>lac</i> (40)	<i>bio</i> (48)	<i>glu</i> (62)
Hfr E — <i>bis</i> (4)	<i>glu</i> (11)	<i>bio</i> (27)	<i>lac</i> (35)	<i>thr</i> (50)

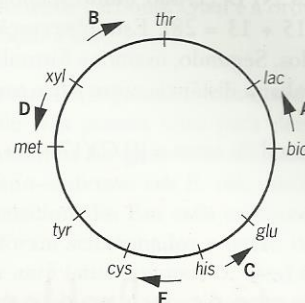
- (a) No mapa a seguir do cromossomo circular de *E. coli*, indique (1) a localização relativa de cada gene, (2) a posição onde o fator sexual está integrado em cada um dos cinco Hfr e (3) o sentido da transferência cromossômica de cada Hfr (indique o sentido com uma seta).



- (b) Para definir melhor a localização da mutação *bis*⁻ no cromossomo, você usa uma linhagem mutante como receptora em um experimento de transdução com bacteriófago P1. Qual, se algum, dos genes mostrado no mapa anteriormente você esperaria que fosse co-transduzido com o alelo *bis*⁺ de seu gene mutante *bis*⁻, considerando que o fago P1 pode acondicionar cerca de 1% da molécula de DNA cromossômico de *E. coli*? Note que o cromossomo de *E. coli* contém 4,6 milhões de pares de nucleotídeos e que a transferência de todo o cromossomo durante a conjugação leva 100 minutos. Explique sua resposta.

Resposta

- (a) A ordem dos genes é a mostrada no mapa a seguir, e os sítios de integração do fator F e sentido de transferência de cada um dos Hfr são indicados por setas marcadas de A a E.



- (b) Nenhum dos marcadores seria co-transduzido com *bis*⁺ porque o fago P1 pode acondicionar apenas 1% do cromossomo de *E. coli*, e nenhum dos outros genes está dentro de 1 minuto de *bis*.

2. Mary Houlahan, George Beadle e Hermione Calhoun (*Genetics* 34:493-507, 1949) estudaram duas linhagens mutantes de *Neurospora crassa*. Uma linhagem (*pdx*) não podia crescer sem piridoxina e a outra (*pan*) não podia crescer sem ácido pantotênico. As tétrades de ascosporos de um cruzamento entre tais linhagens foram dissecadas e analisadas. Pelos resultados (mostrados à direita), determine as relações de ligação entre os genes *pdx* e *pan* e entre cada gene e seu centrômero. (Um + representa o alelo tipo selvagem de um gene; assim, por exemplo, *pdx* + = *pdx pan*⁺ é o genótipo de um esporo que não pode crescer sem piridoxina mas pode crescer sem ácido pantotênico.)

Classes de Tétrades

	1	2	3	4	5	6
<i>pdx</i>	1	<i>pdx pan</i>	<i>pdx</i>	1	<i>pdx</i>	1
<i>pdx</i>	1	<i>pdx pan</i>	<i>pdx</i>	<i>pan</i>	1	<i>pan</i>
1	<i>pan</i>	1	1	1	<i>pdx pan</i>	<i>pdx</i>
1	<i>pan</i>	1	1	<i>pan</i>	1	<i>pan</i>
Número observado	15	1	17	1	13	2

Resposta

Antes de analisar os dados, notamos que o cruzamento foi entre linhagens com os genótipos *pdx* + e + *pan*; estes genótipos são indicativos de ascosporos tipo parental na prole. Os ascosporos não parentais (isto é, recombinantes) são ou *pdx pan* ou + +. Usando esta nomenclatura, podemos classificar cada tétrade como ditipo parental (PD), ditipo não parental (NPD) ou tetratipo (T). Além disso, como as tétrades de *Neurospora* são or-

denadas, podemos classificá-las como dotados de padrões de segregação de primeira (F) ou segunda (S) divisão com relação a cada marcador:

Classe de tétrade	1	2	3	4	5	6
Número observado	15	1	17	1	13	2
Tipo	PD	NPD	T	T	PD	T
Segregação com relação a <i>pdx</i>	F	F	F	S	S	S
Segregação com relação a <i>pan</i>	F	F	S	F	S	S

A análise dos dados envolve várias etapas. Primeira, notamos que as tétrades NPD (1) são muito menos numerosas que as tétrades PD (15 + 13 = 28). Esta observação indica que *pdx* e *pan* estão ligados. Segundo, usamos a fórmula padrão de mapeamento para avaliar a distância entre *pdx* e *pan*:

$$[(1/2)T + NPD]/\text{total de ascas} = [(1/2)(17 + 1 + 2)]/49 = 22,2 \text{ cM}$$

Terceiro, avaliamos a distância entre *pdx* e seu centrômero calculando a frequência dos padrões de segregação da segunda divisão:

$$[(1/2)S]/\text{total de ascas} = [(1/2)(1 + 13 + 2)]/49 = 16 \text{ cM}$$

Similarmente, estimamos a distância entre *pan* e seu centrômero:

$$[(1/2)S]/\text{total de ascas} = [(1/2)(17 + 13 + 2)]/49 = 32,7 \text{ cM}$$

Os resultados combinados nos dizem que *pdx* e *pan* estão do mesmo lado do centrômero e que *pan* está mais longe. Resumimos nossa análise em um mapa de ligação:

$$\text{Centrômero} - 16 \text{ cM} - pdx - 22,2 \text{ cM} - pan$$

Note que a frequência dos padrões de segregação da segunda divisão subestima a distância entre *pan* e o centrômero porque não considera algumas trocas duplas. Podemos obter uma estimativa melhor somando as distâncias entre o centrômero e *pdx* e entre *pdx* e *pan*: $16 + 22,2 = 38,2 \text{ cM}$.

Questões e Problemas

Acentuam a compreensão e desenvolvem as habilidades analíticas.

8.1. Por que critérios os vírus são seres vivos? Não vivos?

8.2. Como os bacteriófagos diferem de outros vírus?

8.3. De que modos os ciclos de vida dos bacteriófagos T4 e λ diferem? Em que aspectos eles são iguais?

8.4. Como a estrutura do profago λ difere da estrutura do cromossomo λ acondicionado no vírion λ ?

8.5. De que modo a integração do cromossomo λ ao cromossomo hospedeiro durante uma infecção lisogênica difere do *crossing-over* entre cromossomos homólogos?

8.6. Como são feitos cruzamentos genéticos com vírus? De que modos esses cruzamentos diferem dos cruzamentos de eucariontes tais como drosófilas?

8.7. O bacteriófago T4 tem um cromossomo linear mas um mapa genético circular. Como as duas estruturas podem ser conciliadas?

8.8. Se um único fago T4, com seu cromossomo linear, infecta uma *E. coli* e passa pelo ciclo lítico, os cromossomos da prole do fago serão circularmente permutados com pontas terminalmente redundantes diferentes. Como são produzidas estas moléculas circularmente permutadas terminalmente redundantes?

8.9. Uma linhagem de *E. coli* nutricionalmente defectiva só cresce em um meio contendo timina, enquanto outra linhagem nutricionalmente defectiva só cresce em meio contendo leucina. Quando essas duas linhagens foram cultivadas juntas,

parte da prole era capaz de crescer em meio mínimo sem timina ou leucina. Como explicar este resultado?

8.10. Suponha que você demonstrou recombinação genética (por ex., quando uma linhagem de genótipo $a^+ b^+$ está presente com uma linhagem de genótipo $a^+ b$, são formados alguns genótipos recombinantes, $a^+ b^+$ e $a b$) em uma espécie antes não estudada de bactérias. Como você determina se a recombinação observada resultou de transformação, conjugação ou transdução?

8.11. Compare as semelhanças e diferenças entre transformação, transdução e sexdução.

8.12. (a) Quais as diferenças genotípicas entre células F^- , F^+ e Hfr ? (b) Quais são as diferenças fenotípicas? (c) Por que mecanismo as células F^- são convertidas em F^+ ? As células F^+ em células Hfr ? As células Hfr em F^+ ?

8.13. (a) Que uso têm os fatores F' na análise genética? (b) Como os fatores F' são formados? (c) Por que mecanismo ocorre sexdução?

8.14. Quais as diferenças básicas entre transdução generalizada e transdução especializada?

8.15. Que papéis os elementos IS desempenham na integração dos fatores F ?

8.16. Como os genes bacterianos podem ser mapeados por experimentos de reprodução interrompida?

8.17. O que significa o termo *co-transdução*? Como as frequências de co-transdução são usadas para mapear marcadores genéticos?

- 8.18. Uma linhagem F^+ , marcada em 10 loci, dá origem espontaneamente a uma prole Hfr sempre que o fator F se torna incorporado ao cromossomo da linhagem F^+ . O fator F pode ser integrado ao cromossomo circular em muitos pontos, de modo que as linhagens Hfr resultantes transferem marcadores genéticos em ordens diferentes. Para qualquer linhagem Hfr, a ordem de marcadores que entram na receptora pode ser determinada por experimentos de reprodução interrompida. A partir dos seguintes dados para várias linhagens Hfr da mesma F^+ , determine a ordem de marcadores na linhagem F^+ .

Linhagem Hfr	Marcadores Doados em Ordem
1	— Z-H-E-R →
2	— O-K-S-R →
3	— K-O-W-I →
4	— Z-T-I-W →
5	— H-Z-T-I →

- 8.19. Em *E. coli*, a capacidade de usar lactose como fonte de carbono requer a presença de enzimas β -galactosidase e β -galactosídeo-permease. Estas enzimas são codificadas por dois genes proximamente ligados, *lacZ* e *lacY*, respectivamente. Outro gene, *proC*, controla, em parte, a capacidade de *E. coli* de sintetizar o aminoácido prolina. Os alelos *str^r* e *str^s*,

respectivamente, controlam resistência e sensibilidade a estreptomicina. Sabe-se que Hfr H transfere os dois genes *lac*, *proC* e *str*, nesta ordem, durante a conjugação.

Foi feito um cruzamento entre Hfr H de genótipo *lacZ⁻ lacY⁺ proC⁺ str^s* e uma linhagem F^- de genótipo *lacZ⁻ lacY⁻ proC⁻ str^r*. Após cerca de 2 horas, a mistura foi diluída e plaqueada em meio contendo estreptomicina mas não prolina. Quando as colônias recombinantes *proC⁺ str^r* resultantes foram examinadas quanto a sua capacidade de crescer em meio contendo lactose como única fonte de carbono, poucas delas foram capazes de fermentar lactose. Quando foi feito o cruzamento recíproco (Hfr H *lacZ⁺ lacY⁻ proC⁺ str^s* X F^- *lacZ⁻ lacY⁺ proC⁻ str^r*), muitos dos recombinantes *proC⁺ str^r* foram capazes de crescer em um meio contendo lactose como única fonte de carbono. Qual é a ordem dos genes *lacZ* e *lacY* em relação a *proC*?

- 8.20. Os dados no quadro seguinte foram obtidos de um teste de transdução de três pontos feito para determinar a ordem dos sítios mutantes no gene *A* codificante da subunidade *a* de triptofano-sintetase em *E. coli*. *Anth* é um marcador ligado não selecionado. Em cada cruzamento, os recombinantes *trp⁺* foram selecionados e então registrados quanto ao marcador *anth* (*anth⁺* ou *anth⁻*). Qual a ordem linear de *anth* e os três alelos mutantes do gene *A* indicados pelos dados no quadro?

Cruzamento	Marcadores de Doadora	Marcadores de Receptora	Alelo <i>anth</i> em Recombinantes	% <i>anth¹</i>
1	<i>anth¹</i> — A34	<i>anth²</i> — A223	72 <i>anth¹</i> : 332 <i>anth²</i>	18
2	<i>anth¹</i> — A46	<i>anth²</i> — A223	196 <i>anth¹</i> : 180 <i>anth²</i>	52
3	<i>anth¹</i> — A223	<i>anth²</i> — A34	380 <i>anth¹</i> : 379 <i>anth²</i>	50
4	<i>anth¹</i> — A223	<i>anth²</i> — A46	60 <i>anth¹</i> : 280 <i>anth²</i>	20

- 8.21. Duas mutações adicionais no gene *trp A* de *E. coli*, *trp A58* e *trp A487*, foram ordenadas com relação a *trp A223* e ao marcador externo *anth* por cruzamentos de transdução de três fatores

conforme descrito no Problema 8.20. Os resultados destes cruzamentos estão resumidos no quadro seguinte. Qual a ordem linear de *anth* e dos três sítios mutantes no gene *trp A*?

Cruzamento	Marcadores de Doadora	Marcadores de Receptora	Alelo <i>anth</i> em Recombinantes <i>trp¹</i>	% <i>anth²</i>
1	<i>anth¹</i> — A487	<i>anth²</i> — A223	72 <i>anth¹</i> : 332 <i>anth²</i>	82
2	<i>anth¹</i> — A58	<i>anth²</i> — A223	196 <i>anth¹</i> : 180 <i>anth²</i>	48
3	<i>anth¹</i> — A223	<i>anth²</i> — A487	380 <i>anth¹</i> : 379 <i>anth²</i>	50
4	<i>anth¹</i> — A223	<i>anth²</i> — A58	60 <i>anth¹</i> : 280 <i>anth²</i>	80

- 8.22. Sabe-se que as mutações *nrd 11* (gene *nrd B*, que codifica a subunidade beta da enzima ribonucleotídeo-redutase), *am M69* (gene 63, que codifica uma proteína auxiliadora da ligação da fibra da cauda) e *nd 28* (*denA*, que codifica a enzima endonuclease II) estão situadas entre o gene 31 e o

gene 32 no cromossomo do bacteriófago T4. As mutações *am N54* e *am A453* estão situadas nos genes 31 e 32, respectivamente. Considerando os dados no cruzamento de três fatores do quadro seguinte, qual a ordem linear dos cinco sítios mutantes?

Dados de Cruzamento de Três Fatores			% Recombinação*
1.	<i>am</i> A453— <i>am</i> M69 X <i>nrd</i> 11		2,6
2.	<i>am</i> A453— <i>nrd</i> 11 X <i>am</i> M69		4,2
3.	<i>am</i> A453— <i>am</i> M69 X <i>nd</i> 28		2,5
4.	<i>am</i> A453— <i>nd</i> 28 X <i>am</i> M69		3,5
5.	<i>am</i> A453— <i>nrd</i> 11 X <i>nd</i> 28		2,9
6.	<i>am</i> A453— <i>nd</i> 28 X <i>nrd</i> 11		2,1
7.	<i>am</i> N54— <i>am</i> M69 X <i>nrd</i> 11		3,5
8.	<i>am</i> N54— <i>nrd</i> 11 X <i>am</i> M69		1,9
9.	<i>am</i> N54— <i>nd</i> 28 X <i>am</i> M69		1,7
10.	<i>am</i> N54— <i>am</i> M69 X <i>nd</i> 28		2,7
11.	<i>am</i> N54— <i>nd</i> 28 X <i>nrd</i> 11		2,9
12.	<i>am</i> N54— <i>nrd</i> 11 X <i>nd</i> 28		1,9

*Todas as frequências de recombinação são dadas como

$$\frac{2 \text{ (prole tipo selvagem)}}{\text{prole total}} \times 100.$$

8.23. Em leveduras, a mutação *ad* requer adenina para crescimento e a mutação *in* requer inositol para crescimento. Os alelos tipo selvagem destas duas mutações são *AD* e *IN*. C. C. Lindgren analisou 48 tetrades do cruzamento *AD IN* × *ad in*; entre essas tetrades, 22 eram ditipo parental, 3 eram ditipo não parental e 23 eram tetratipo. Os genes *AD* e *IN* são ligados? Caso sim, qual a distância entre eles?

8.24. Um geneticista obteve os seguintes dados de tetrade ordenada de um cruzamento em *Neurospora*:

Pares de Esporos				Número de Tétrades
Topo do Asco	Inferior do Asco			
(1,2)	(3,4)	(5,6)	(7,8)	
<i>A</i>	<i>A</i>	<i>a</i>	<i>a</i>	61
<i>a</i>	<i>a</i>	<i>A</i>	<i>A</i>	55
<i>a</i>	<i>A</i>	<i>a</i>	<i>A</i>	40
<i>A</i>	<i>a</i>	<i>A</i>	<i>a</i>	44
Total 5				200

Qual a distância entre o locus *A* e o centrômero no cromossomo?

8.25. Os dados da seguinte tetrade foram obtidos do cruzamento *AB* × *a b* em *Neurospora*:

Pares de Esporos				Número de Tétrades
Topo do Asco	Inferior do Asco			
(1,2)	(3,4)	(5,6)	(7,8)	
<i>AB</i>	<i>AB</i>	<i>ab</i>	<i>ab</i>	1.766
<i>AB</i>	<i>aB</i>	<i>Ab</i>	<i>ab</i>	220
<i>AB</i>	<i>Ab</i>	<i>aB</i>	<i>ab</i>	14
Total 5				2.000

Use estes dados para construir um mapa de ligação dos dois genes e o centrômero em seu cromossomo.

8.26. As análises de tetrades desordenadas de leveduras do cruzamento +++ × *a b c* produziram os seguintes dados:

Classe de Tetrade	Esporos				Número de Ascós
1.	<i>a b c</i>	<i>a b c</i>	1 1 1	1 1 1	36
2.	<i>a b c</i>	<i>a 1 c</i>	1 <i>b</i> 1	1 1 1	14
3.	<i>a 1 1</i>	<i>a 1 1</i>	1 <i>b c</i>	1 <i>b c</i>	32
4.	<i>a b 1</i>	<i>a 1 1</i>	1 <i>b c</i>	1 1 <i>c</i>	16
5.	<i>a b 1</i>	<i>a b 1</i>	1 1 <i>c</i>	1 1 <i>c</i>	2

Quais destes genes são ligados e qual a distância de mapa entre eles?

8.27. Em *Neurospora*, as mutações *arg*, *thi* e *leu* bloqueiam a síntese de arginina, tiamina e leucina, respectivamente. Os seguintes dados de tetrade ordenada vieram de dois cruzamentos com estas mutações:

Cruzamento	Pares de Esporos				Número de Tétrades
	(1,2)	(3,4)	(5,6)	(7,8)	
<i>arg</i> 3 <i>thi</i>	<i>arg</i> 1	<i>arg</i> 1	1 <i>thi</i>	1 <i>thi</i>	46
	<i>arg thi</i>	<i>arg thi</i>	1 1	1 1	56
Total 5					100
<i>arg</i> 3 <i>leu</i>	<i>arg</i> 1	<i>arg</i> 1	1 <i>leu</i>	1 <i>leu</i>	155
	<i>arg</i> 1	<i>arg leu</i>	1 1	1 <i>leu</i>	44
	<i>arg</i> 1	1 <i>leu</i>	<i>arg</i> 1	1 <i>leu</i>	1
Total 5					200

Usando estes dados, construa o(s) mapa(s) de ligação para estes genes. Mostre a posição do(s) centrômero(s).

8.28. A ordem de três genes e o centrômero em um cromossomo de *Neurospora* é centrômero — *x* — *y* — *z*. Um cruzamento ente +++ e *x y z* produziu um asco com o seguinte arranjo ordenado de ascósporos (apenas um membro de cada par de esporos é mostrado): (+ + *z*) (+ *y z*) (*x* + +) (*x y* +).

- Este asco é mais provavelmente o resultado de um evento meiótico no qual ocorreu 0, 1, 2 ou 3 *crossings*?
- Em que intervalo(s) o(s) *crossing*(s) mais provavelmente ocorreu(eram)?
- Se *crossings* duplos ou triplos estavam envolvidos, eles foram de dois filamentos, três ou *crossings* múltiplos de quatro filamentos?

8.29. De um cruzamento de linhagem de *Neurospora* tipo selvagem com uma linhagem que tinha necessidades nutricionais dos aminoácidos histidina (*his*) e tirosina (*tir*), foram obtidas as seguintes tetrades ordenadas:

- Qual a distância de mapa entre o locus *his* e o centrômero?
- Qual a relação de ligação entre os loci *his* e *tir*?

Pares de Esporos	Tétrades							
	1	2	3	4	5	6	7	8
1-2	<i>bis tyr</i>	<i>bis tyr</i>	<i>bis tyr</i>	<i>bis 1</i>	<i>bis 1</i>	1 <i>tyr</i>	1 <i>tyr</i>	<i>bis 1</i>
3-4	<i>bis tyr</i>	<i>bis 1</i>	1 <i>tyr</i>	<i>bis 1</i>	<i>bis tyr</i>	<i>bis tyr</i>	1 1	1 <i>tyr</i>
5-6	1 1	1 <i>tyr</i>	1 1	1 <i>tyr</i>	1 <i>tyr</i>	<i>bis 1</i>	<i>bis tyr</i>	<i>bis tyr</i>
7-8	1 1	1 1	<i>bis 1</i>	1 <i>tyr</i>	1 1	1 1	<i>bis 1</i>	1 1
Número	30	8	5	30	10	7	8	2
Total 5								100

8.30. As seguintes tétrades foram produzidas por um cruzamento de uma *Neurospora* que tinha esporos brancos (*w*) e uma necessidade nutricional do aminoácido arginina (*arg*) com uma linhagem que tinha esporos escuros e nenhuma necessidade de arginina:

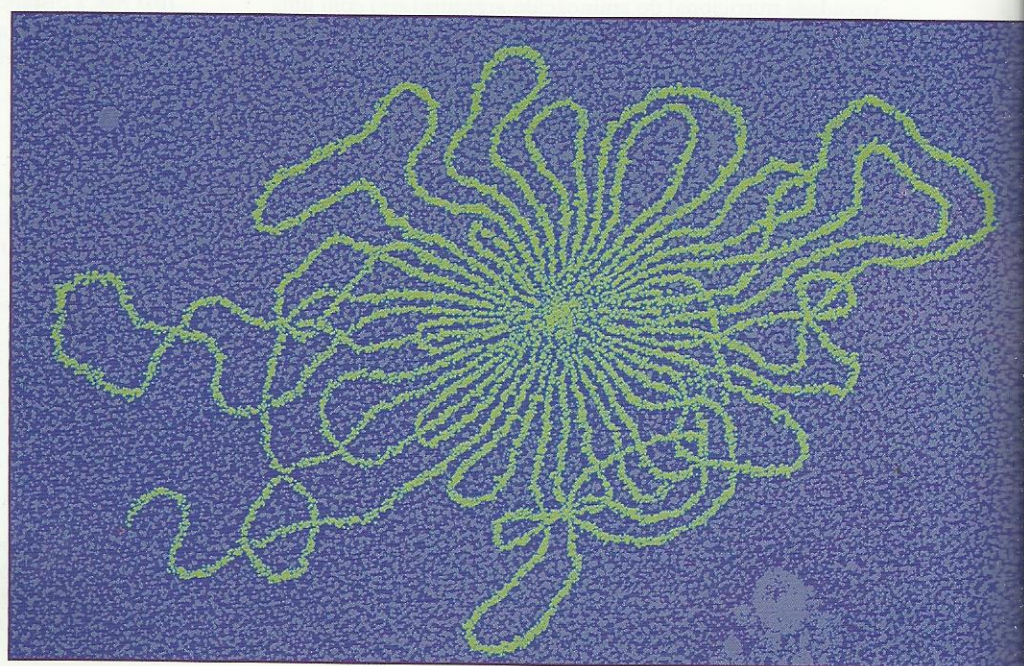
Pares de Esporos	Tétrades					
	1	2	3	4	5	6
1-2	<i>w arg</i>	<i>w arg</i>	<i>w arg</i>	<i>w arg</i>	<i>w +</i>	<i>w +</i>
3-4	<i>w arg</i>	<i>w +</i>	<i>+ arg</i>	<i>++</i>	<i>w +</i>	<i>++</i>
5-6	<i>++</i>	<i>+ arg</i>	<i>++</i>	<i>+ arg</i>	<i>+ arg</i>	<i>w arg</i>
7-8	<i>++</i>	<i>++</i>	<i>w +</i>	<i>w +</i>	<i>+ arg</i>	<i>+ arg</i>
Número	58	14	15	2	1	10
Total =						100

- Qual a distância de mapa entre o *locus arg* e o centrômero?
- Qual a distância de mapa entre os loci *w* e *arg*?
- Os loci *arg* e *w* são encontrados no mesmo braço ou em braços diferentes do cromossomo?

DNA e Estrutura Molecular de Cromossomos

Tópicos

- Funções do Material Genético
- Prova de que a Informação Genética É Estocada no DNA
- As Estruturas do DNA e do RNA
- Estrutura Cromossômica em Procariontes e Vírus
- Estrutura Cromossômica em Eucariontes
- Genomas Eucarióticos: Sequências Repetidas de DNA



Micrografia eletrônica de transmissão com cor acentuada do DNA do bacteriófago lambda.

Descoberta da Nucleína

Em 1868, Johann Friedrich Miescher, um jovem estudante de medicina suíço, ficou fascinado com uma substância ácida que ele isolou de células de pus obtidas de bandagens usadas para cobrir feridas humanas. Primeiro ele separou as células do pus das bandagens e restos associados e então tratou as células com pepsina, uma enzima proteolítica que ele isolou dos estômagos de porcos. Após o tratamento com pepsina, ele recuperou uma substância ácida que ele chamou de “nucleína”. A nucleína de Miescher era incomum por conter grandes quantidades de nitrogênio e fósforo, dois elementos que se sabia na época coexistir apenas em alguns tipos de gordura. Miescher publicou um trabalho descrevendo sua descoberta da nucleína em células de pus humano e o submeteu para publicação em 1869. Entretanto, o editor do periódico ao qual o trabalho foi enviado desconfiou dos resultados e decidiu ele mesmo repetir os experimentos. Como consequência, o trabalho

de Miescher descrevendo a nucleína só foi publicado em 1871, dois anos após sua apresentação.

Na época, a importância da substância que Miescher chamou de nucleína não podia ser antecipada. A existência de cadeias polinucleotídicas, o componente principal do material ácido na nucleína de Miescher, só foi documentada nos anos 1940. O papel dos ácidos nucleicos em estocar e transmitir informações genéticas só foi estabelecido em 1944, e a estrutura da dupla hélice do DNA só foi descoberta em 1953. Mesmo em 1953, muitos geneticistas relutaram em aceitar a idéia de que os ácidos nucleicos, e não as proteínas, levavam a informação genética porque os ácidos nucleicos apresentavam menos variabilidade estrutural que as proteínas.

FUNÇÕES DO MATERIAL GENÉTICO

O material genético deve replicar-se, controlar o crescimento e o desenvolvimento do organismo e permitir que o organismo se adapte a mudanças no ambiente.

Em 1865, Mendel mostrou que “*Merkmale*” (hoje “genes”) transmitem a informação genética, e, na primeira parte do século vinte, foram amplamente estudados seus padrões de transmissão. Embora tais estudos clássicos de genética dessem pouca elucidação quanto à natureza molecular dos genes, eles demonstraram que o material genético devia desempenhar três funções essenciais:

1. A função genotípica, a **replicação**. O material genético deve estocar a informação genética e transmitir com precisão essa informação dos genitores para a prole, geração após geração.
2. A função fenotípica, a **expressão gênica**. O material genético deve controlar o desenvolvimento do fenótipo do organismo. Isto é, o material genético deve ditar o crescimento do organismo desde o zigoto unicelular até o adulto.
3. A função evolutiva, as **mutações**. O material genético deve sofrer mudanças para produzir variações que permitem aos organismos adaptar-se a modificações no ambiente de modo que possa ocorrer evolução.

Outros estudos genéticos iniciais estabeleceram uma correlação precisa entre os padrões de transmissão dos genes e o comportamento dos cromossomos durante a reprodução sexual, dando fortes evidências de que os genes geralmente estão situados em cromossomos. Assim, tentativas posteriores de descobrir a base química da hereditariedade enfocaram moléculas presentes nos cromossomos.

Os cromossomos são compostos de dois tipos de grandes moléculas orgânicas (macromoléculas) chamadas **proteínas** e **ácidos nucleicos**. Os ácidos nucleicos são de dois tipos: o **ácido desoxirribonucleico (DNA)** e o **ácido ribonucleico (RNA)**. Durante a década de 1940 e o começo da de 1950,

os resultados de elegantes experimentos estabeleceram claramente que a informação genética é estocada em ácidos nucleicos, e não em proteínas. Na maioria dos organismos, a informação genética é codificada na estrutura do DNA. Entretanto, em muitos vírus pequenos, a informação genética é codificada no RNA.

PONTOS IMPORTANTES

- O material genético deve desempenhar três funções essenciais: a função genotípica – replicação –, a função fenotípica – expressão gênica – e a função evolutiva – mutação.

PROVA DE QUE A INFORMAÇÃO GENÉTICA É ESTOCADA NO DNA

Na maioria dos organismos, a informação genética está codificada no DNA. Em alguns vírus, o RNA é o material genético. Viróides são moléculas infecciosas de RNA nu, e prions são proteínas infecciosas herdáveis.

Várias linhas de evidências indiretas sugeriram que o DNA abriga a informação genética de organismos vivos. Por exemplo, a maior parte do DNA das células está situada nos cromossomos, enquanto RNA e proteínas são também abundantes no citoplasma. Além disso, existe uma correlação precisa entre a quantidade de DNA por célula e o número de conjuntos de cromossomos por célula. A maioria das células somáticas de organismos diplóides contém duas vezes a quantidade de DNA das células germinativas (gametas) da mesma espécie. A composição molecular do DNA é a mesma (com raras exceções) em todas as células de um organismo, enquanto a composição tanto do RNA quanto de proteínas é altamente variável de um tipo de célula para outro. O DNA é mais estável que o RNA ou as proteínas. Como o material genético deve estocar e transmitir informações dos genitores para a prole, podemos esperar que ele seja estável, como o DNA. Embora tais correlações sugiram fortemente que o DNA é o material genético, elas não o comprovam.

Prova de que o DNA Medeia a Transformação

A descoberta de Frederick Griffith da transformação em *Streptococcus pneumoniae* foi discutida no Cap. 8. Quando Griffith injetou tanto bactérias tipo IIIS mortas por aquecimento (virulentas quando vivas) e bactérias tipo IIR vivas (não virulentas) em camundongos, muitos dos camundongos desenvolveram pneumonia e morreram, e células tipo IIIS vivas foram recuperadas de seus cadáveres. Algo nas células mortas pelo calor, o “princípio transformante”, tinha convertido as células tipo IIR vivas em tipo IIIS. Em 1931, Richard Sia e Martin Dawson fizeram o mesmo experimento *in vitro*, mostrando que o camundongo não teve nenhum papel no processo de transformação (Fig. 9.1). O experimento de Sia e Dawson abriu o caminho para a demonstração de Oswald Avery, Colin MacLeod e Maclyn McCarty de que o princípio transformante em *S. pneumoniae* é o DNA. Avery e colaboradores mostraram que o DNA é o único componente das células IIIS necessário para transformar o tipo IIR em tipo IIIS (Fig. 9.2).

Mas como eles podiam estar certos de que o DNA estava de fato puro? Provar a pureza de qualquer substância macromolecular é extremamente difícil. Talvez a preparação de

DNA contivesse algumas moléculas de proteína, e tais proteínas contaminantes fossem responsáveis pela transformação observada. Os experimentos mais definitivos de Avery, MacLeod e McCarty provando que o DNA era o princípio transformante envolveram o uso de enzimas que degradam DNA, RNA ou proteínas. Em experimentos separados, o DNA altamente purificado de células tipo IIIS foi tratado com as enzimas (1) **desoxirribonuclease (DNase)**, que degrada DNA, (2) **ribonuclease (RNase)**, que degrada RNA ou (3) **proteases**, que degradam proteínas. O DNA foi então testado quanto a sua capacidade de transformar células tipo IIR em tipo IIIS. Apenas o tratamento com DNase tinha qualquer efeito sobre a atividade transformante da preparação de DNA – ele eliminava toda a atividade transformante (Fig. 9.2).

Embora o mecanismo molecular pelo qual ocorre a transformação tivesse ficado desconhecido por muitos anos, os resultados de Avery e colaboradores estabeleceram claramente que a informação genética em *Streptococcus* está presente no DNA. Os geneticistas hoje sabem que o segmento de DNA no cromossomo de *Streptococcus* portador da informação genética que especifica a síntese de uma cápsula tipo III é fisicamente inserido no cromossomo da receptora tipo IIR durante o processo de transformação.

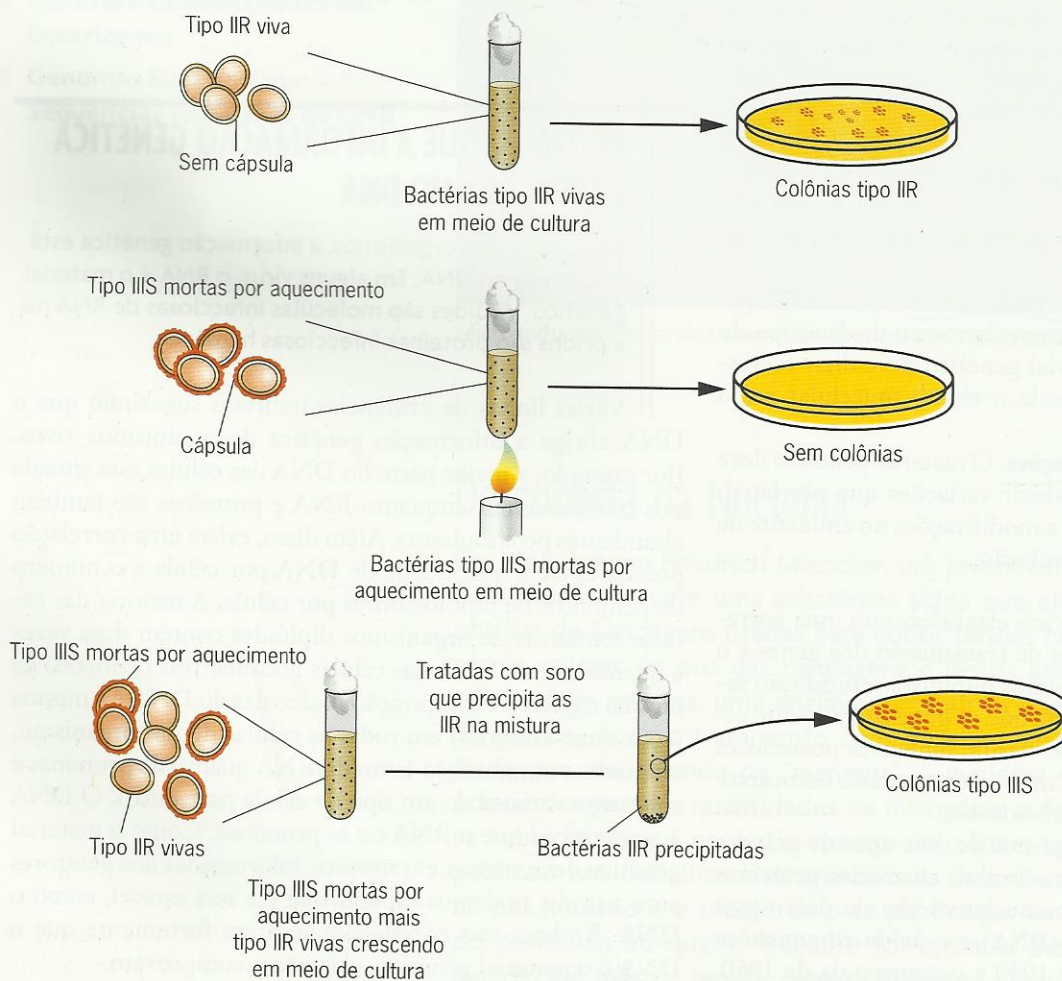


Fig. 9.1 ■ Demonstração de Sia e Dawson de transformação em *Streptococcus pneumoniae* *in vitro*.

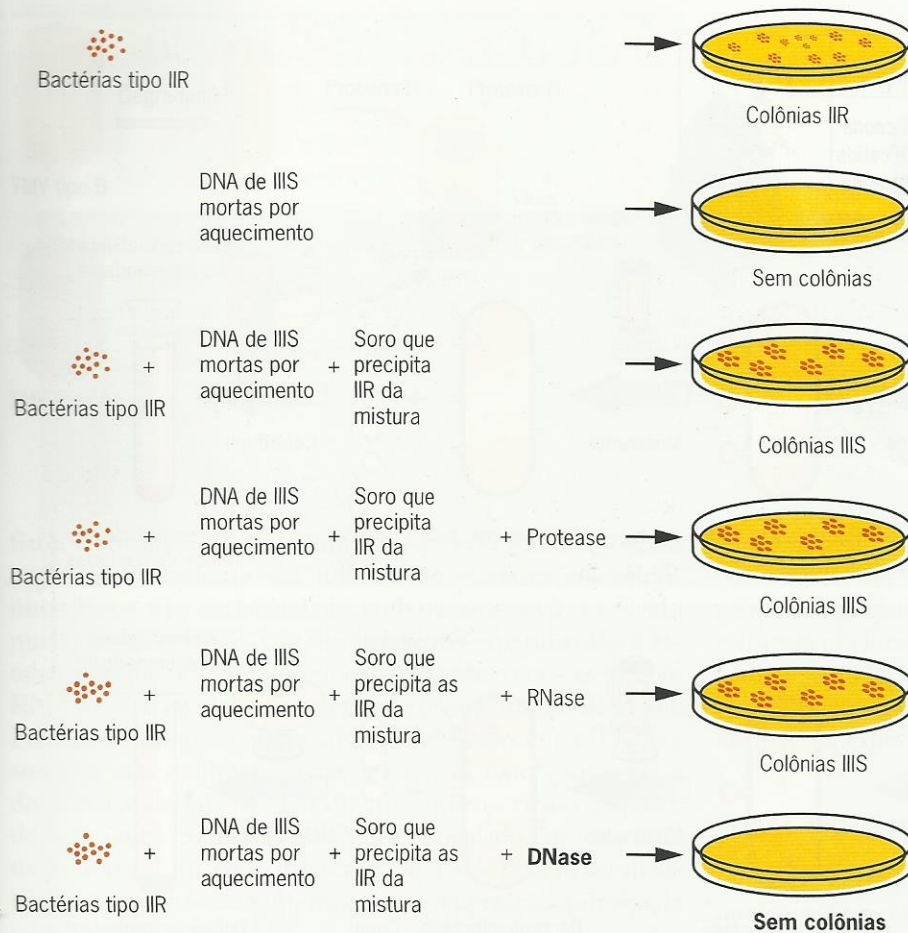


Fig. 9.2 ■ A prova de Avery, MacLeod e McCarty de que o “princípio transformante” é o DNA.

Prova de que o DNA Leva a Informação Genética no Bacteriófago T2

Uma evidência adicional demonstrando que o DNA é o material genético foi publicada em 1952 por Alfred Hershey (ganhador do Prêmio Nobel de 1969) e Martha Chase. Os resultados de seus experimentos mostraram que a informação genética de um determinado vírus de bactérias (bacteriófago T2) estava presente em seu DNA. Seus resultados tiveram um grande impacto sobre a aceitação pelos cientistas de que o DNA é o material genético. Esse impacto foi o resultado da simplicidade do experimento de Hershey-Chase.

Vírus são os menores organismos vivos. Eles estão vivos pelo menos no sentido de que sua reprodução é controlada pela informação genética estocada em ácidos nucleicos mediante os mesmos processos de organismos celulares (Cap. 8). Entretanto, os vírus são parasitas acelulares que podem reproduzir-se apenas em células hospedeiras apropriadas. Sua reprodução é totalmente dependente da maquinaria metabólica (ribossomos, sistemas geradores de energia e outros componentes) do hospedeiro. Os vírus foram extremamente úteis no estudo de muitos processos genéticos por causa de sua estrutura e composição química simples (muitos contêm apenas proteínas e ácidos nucleicos) e de sua rápida repro-

dução (15 a 20 minutos em caso de alguns vírus bacterianos em condições ótimas).

O bacteriófago T2, que infecta o bacilo intestinal comum *Escherichia coli*, é composto de cerca de 50% de DNA e cerca de 50% de proteína (Fig. 9.3). Os experimentos anteriores a 1952 haviam mostrado que toda a reprodução do bacteriófago T2 ocorria dentro de *E. coli*. Portanto, quando Hershey e Chase mostraram que o DNA da partícula viral entrava na célula, enquanto a maior parte da proteína do vírus era adsorvida para o exterior da célula, ficou claro que a informação genética necessária para reprodução viral estava presente no DNA. A base para o experimento de Hershey-Chase é que o DNA contém fósforo mas não enxofre, enquanto as proteínas contêm enxofre mas não fósforo. Assim, Hershey e Chase foram capazes de marcar especificamente (1) o DNA do fago mediante crescimento em um meio contendo o isótopo radioativo do fósforo, ^{32}P , em lugar do isótopo normal, ^{31}P ; ou (2) a capa de proteína do fago mediante crescimento em um meio contendo enxofre radioativo, ^{35}S , em lugar do isótopo normal, ^{32}S (Fig. 9.3).

Quando partículas de fago T2 marcadas com ^{35}S foram misturadas com *E. coli* por alguns minutos e as células infectadas pelo fago foram então submetidas a forças de separação em um misturador, a maior parte da radioatividade (e por-

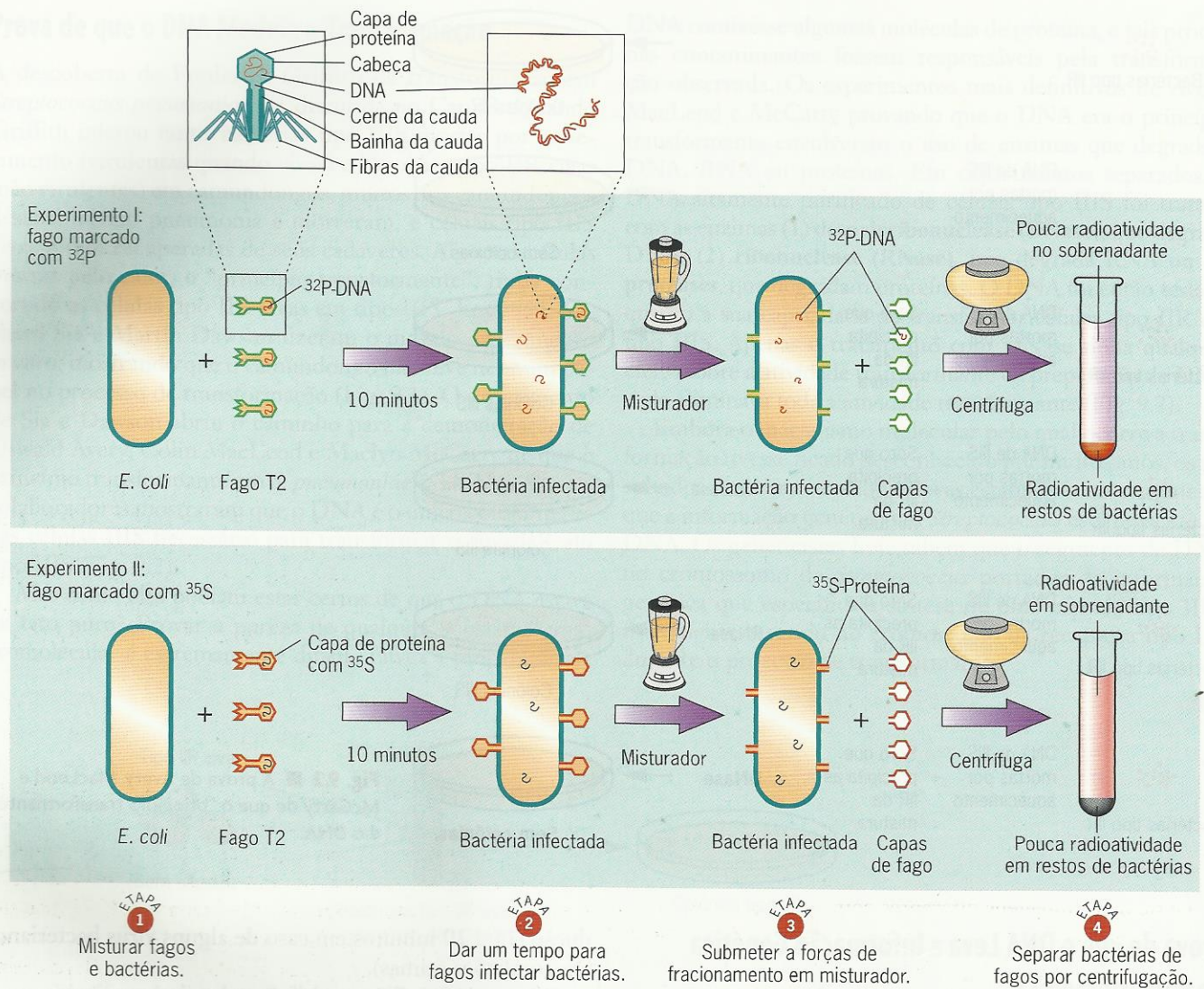


Fig. 9.3 ■ Demonstração por Hershey e Chase de que a informação genética do bacteriófago T2 está em seu DNA.

tanto as proteínas) podia ser removida das células sem afetar a produção da prole de fagos. Quando foram usadas partículas T2 nas quais o DNA foi marcado com ^{32}P , entretanto, toda a radioatividade foi encontrada dentro das células; isto é, o DNA não foi removido por agitação no misturador. As capas removidas dos fagos foram separadas das células infectadas por centrifugação em baixa velocidade, com sedimentos de células deixando as partículas de fago suspensas. Tais resultados indicaram que o DNA do vírus entrava na célula hospedeira, enquanto a capa de proteína ficava do lado de fora da célula. Como a prole dos vírus é produzida dentro das células, os resultados de Hershey e Chase indicaram que a informação genética que dirige a síntese tanto de moléculas de DNA quanto das capas de proteína da prole de vírus deve estar presente no DNA parental. Além disso, as partículas da prole continham parte do ^{32}P , mas nada do ^{35}S do fago parental.

Havia uma falha na prova de Hershey e Chase de que o material genético do fago T2 era o DNA. Verificou-se que

uma quantidade significativa de ^{35}S (e portanto proteína) foi injetada nas células hospedeiras com o DNA. Assim, poderia ser questionado que essa pequena fração de proteínas do fago continha a informação genética. Mais recentemente, os cientistas desenvolveram procedimentos pelos quais protoplastos (células com as paredes removidas) de *E. coli* podem ser infectados com DNA puro do fago. Proles normais de fagos infectivos são produzidas nesses experimentos, chamados experimentos de **transfecção**, provando que o material genético de tais vírus bacterianos é o DNA.

Prova de que o RNA Estoca a Informação Genética em Alguns Vírus

À medida que mais vírus foram identificados e estudados, tornou-se aparente que muitos deles continham RNA e proteínas mas não DNA. Em todos os casos estudados até agora,

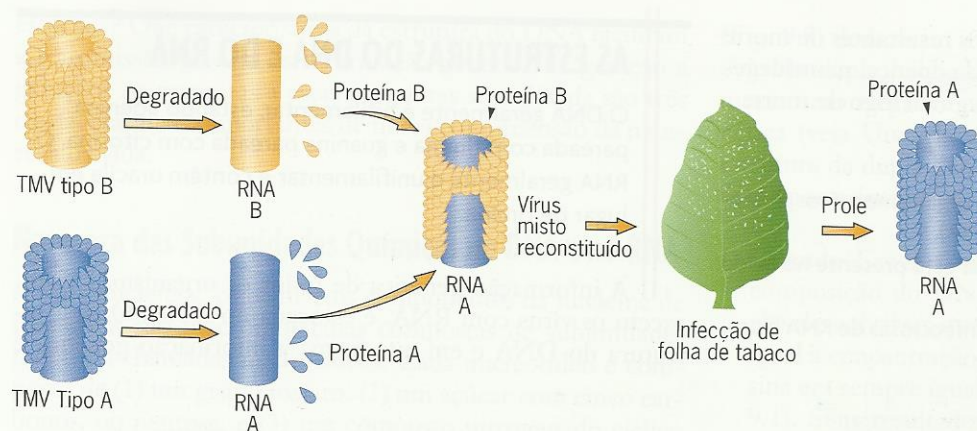


Fig. 9.4 ■ Demonstração de que o material genético do vírus do mosaico do tabaco (TMV) é o RNA, e não proteína. O TMV não contém DNA; ele é composto apenas de RNA e proteína.

ficou claro que tais vírus com RNA, como todos os outros organismos, estocam sua informação genética em ácidos nucleicos e não em proteínas, embora nesses vírus o ácido nucleico seja o RNA. Um dos primeiros experimentos a estabelecer que o RNA era o material genético nos vírus com RNA foi o chamado experimento de reconstituição de Heinz Fraenkel-Conrat e colaboradores, publicado em 1957. Seu simples, mas definitivo, experimento foi feito com o vírus do mosaico do tabaco (TMV), um pequeno vírus composto de uma única molécula de RNA encapsulado em uma capa de proteína. Linhagens diferentes de TMV podem ser identificadas com base em diferenças na composição química de suas capas de proteína.

Fraenkel-Conrat e colaboradores trataram partículas de TMV de duas linhagens diferentes com substâncias químicas que dissociam as capas de proteína dos vírus das moléculas de RNA e separam as proteínas do RNA. Eles então misturaram as proteínas de uma linhagem com as moléculas de RNA da outra linhagem em condições que resultam na reconstituição de vírus infecciosos completos compostos de proteínas de uma linhagem e RNA de outra linhagem. Quando as folhas de tabaco foram infectadas com tais vírus mistos reconstituídos, os vírus da prole eram sempre fenotipicamente e genotipicamente idênticos à linhagem parental da qual tinha sido obtido o RNA (Fig. 9.4). Assim, a informação genética do TMV está estocada no RNA, e não em proteína.

Viróides, Moléculas Nuas Infecciosas de RNA Herdáveis

Alguns agentes infecciosos que causam doenças em plantas, e outros que são tidos como causadores de doenças em animais contêm pequenas moléculas circulares de RNA. Entretanto, ao contrário dos vírus com RNA tais como o TMV, essas moléculas de RNA não são acondicionadas em capas de proteína. Elas foram descobertas por Theodore Diener e colaboradores nos anos 1960 e denominadas **viróides**, significando “similares a vírus”. O viróide da batata (PSTV) foi um dos primeiros viróides estudados. Ele faz com que os tubérculos das batatas infectadas sejam longos e pontiagudos

como hastes. O PSTV é uma molécula circular de RNA que consiste em 359 nucleotídeos. Ele se replica autonomamente em células do tubérculo da batata. O PSTV tem uma estrutura em forma de bastão devido ao pareamento de bases dentro da molécula. Os efeitos patogênicos dos viróides parecem ser resultantes de sua capacidade de alterar os padrões normais da expressão gênica.

Prions, Proteínas Infecciosas Herdáveis

Outros agentes infecciosos transmissíveis não contêm ácido nucleico de nenhum tipo, apenas proteína. Esses agentes proteináceos foram chamados **prions** (derivados de *proteína* e *infecciosa*) em 1982 por Stanley Prusiner, que recebeu o Prêmio Nobel de Fisiologia ou Medicina em 1997 por sua pesquisa sobre estas proteínas únicas. Os prions são formas alteradas de proteínas celulares normais em mamíferos. Eles são responsáveis por um grupo de doenças neurodegenerativas fatais incluindo doença Creutzfeldt-Jakob (CJD) e kuru em humanos, encefalopatia espongiforme bovina (BSE), comumente chamada “doença da vaca louca” em gado, doença do definhamento crônico (CWD) em cervos e alces e *scrapie* em ovelhas. A importância destas doenças foi documentada pela mortalidade de rebanhos inteiros de gado expostos a BSE no Reino Unido e alces expostos a CWD nos EUA durante a última década. As doenças de prion são espalhadas pelo consumo de carne contendo as proteínas infecciosas. No caso de kuru nos habitantes de Fore na Nova Guiné, acredita-se que a doença seja transmitida pelo canibalismo ritualístico dos cérebros de parentes mortos. Sabe-se que o “mal da vaca louca” se dissemina quando tecidos de animais infectados são usados em suplementos alimentares de gado.

Os prions são formas aberrantes de proteínas codificadas por genes normais de mamíferos. Quando a forma aberrante patogênica da proteína, o prion, é formada, ela atua como um molde que converte mais proteínas celulares normais para a forma infecciosa de prion. As proteínas prion aberrantes aglomeram-se e acabam por matar as células hospedeiras. Os cérebros dos animais doentes desenvolvem uma morfo-

logia espongiiforme com espaços vazios resultantes de morte celular. Quando surgem os sintomas da doença, neurodegeneração ocorre muito rapidamente, seguida logo de morte.

PONTOS IMPORTANTES

- A informação genética da maioria dos organismos vivos é estocada no ácido desoxirribonucleico (DNA).
- Em alguns vírus, a informação genética está presente no ácido ribonucleico (RNA).
- Viróides e prions são moléculas nuas infecciosas de RNA e proteína, respectivamente.

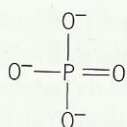
AS ESTRUTURAS DO DNA E DO RNA

O DNA geralmente é bifilamentar, estando adenina pareada com timina e guanina pareada com citosina. O RNA geralmente é unifilamentar e contém uracila em lugar de timina.

A informação genética de todos os organismos vivos, exceto os vírus com RNA, é estocada no DNA. Qual a estrutura do DNA e em que forma a informação genética é

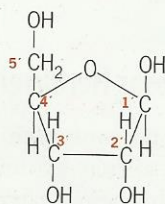
Ácidos nucleicos são compostos de subunidades repetidas chamadas nucleotídeos. Cada nucleotídeo é composto de três unidades.

(1)
Um
grupo
fosfato:

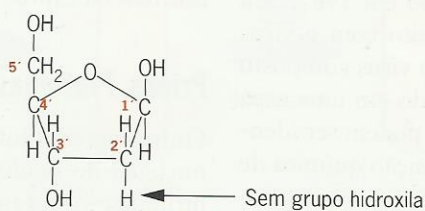


(2)
Um
açúcar de
cinco carbonos
ou pentose:

(a) No RNA:
Ribose

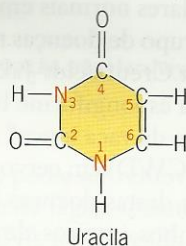


(b) No DNA:
2-Desoxirribose

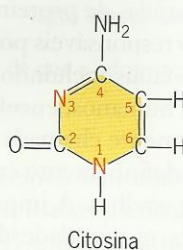


(3)
Uma
base
cíclica
contendo
nitrogênio:

(a) No RNA apenas
(com raras exceções):



(b) No RNA e
no DNA:



(c) Só no DNA
(com raras exceções):

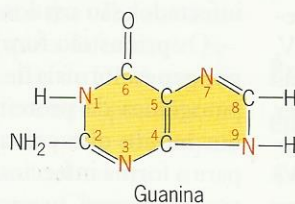
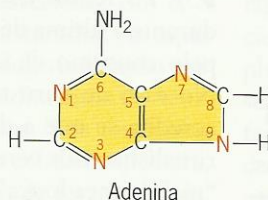
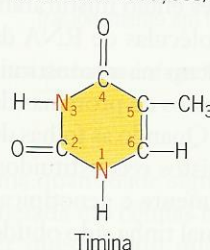


Fig. 9.5 ■ Componentes estruturais dos ácidos nucleicos.

estocada? Que características da estrutura do DNA facilitam a transmissão precisa da informação genética de geração a geração? As respostas a estas perguntas sem dúvida são três das facetas mais importantes de nossa compreensão da natureza da vida.

Natureza das Subunidades Químicas no DNA e no RNA

Ácidos nucleicos, os principais componentes da nucleína de Miescher, são macromoléculas compostas de subunidades repetidas chamadas **nucleotídeos**. Cada nucleotídeo é composto de (1) um grupo fosfato, (2) um açúcar com cinco carbonos, ou pentose, e (3) um composto nitrogenado cíclico chamado base (Fig. 9.5). No DNA, o açúcar é a 2-desoxirribose (donde o nome ácido desoxirribonucleico). No RNA, o açúcar é a ribose (donde ácido ribonucleico). Quatro bases diferentes são comumente encontradas no DNA: **adenina (A)**, **guanina (G)**, **timina (T)** e **citossina (C)**. O RNA também contém adenina, guanina e citossina, mas tem uma base diferente, **uracila (U)**, em lugar de timina. Adenina e guanina são bases com dois anéis chamadas **purinas**; citossina, timina e uracila são bases com um só anel chamadas **pirimidinas**. Tanto o DNA quanto o RNA, portanto, contêm quatro subunidades diferentes, ou nucleotídeos: dois nucleotídeos purina e dois nucleotídeos pirimidina (Fig. 9.6). Em polinucleotídeos tais como DNA e RNA, essas subunidades são unidas em longas cadeias (Fig. 9.7). O RNA geralmente existe como um polímero unifilamentar que é composto de uma longa sequência de nucleotídeos. O DNA tem um nível de organização adicional e muito importante: geralmente é uma molécula bifilamentar.

Estrutura do DNA: A Dupla Hélice

Uma das mais excitantes descobertas na história da biologia ocorreu em 1953 quando James Watson e Francis Crick

(Fig. 9.8) deduziram a estrutura correta do DNA. Seu modelo da dupla hélice do DNA imediatamente sugeriu um mecanismo elegante para a transmissão da informação genética (veja Um Marco na Genética: A Dupla Hélice). A estrutura da dupla hélice de Watson e Crick foi baseada em dois tipos principais de evidência:

1. Quando Erwin Chargaff e colaboradores analisaram a composição do DNA de muitos organismos diferentes, eles descobriram que a concentração de timina era sempre igual à concentração de adenina e a concentração de citossina era sempre igual à concentração de guanina (Quadro 9.1). Seus resultados sugerem fortemente que timina e adenina bem como citossina e guanina estão presentes no DNA em alguma inter-relação fixa. Seus dados também mostraram que a concentração total de pirimidinas (timina mais citossina) era sempre igual à concentração total de purinas (adenina mais guanina; veja Quadro 9.1).
2. Quando raios X passam pelas fibras de moléculas purificadas, eles são desviados pelos átomos das moléculas de modos específicos, chamados padrões de difração, que dão informações sobre a organização dos componentes das moléculas. Esses **padrões de difração de raios X** podem ser registrados em filmes sensíveis aos raios X do mesmo modo que os padrões de luz podem ser registrados com uma câmara e um filme sensível a luz. Watson e Crick usaram os dados de difração de raios X na estrutura do DNA (Fig. 9.9) fornecidos por Maurice Wilkins, Rosalind Franklin e seus colaboradores. Esses dados indicaram que o DNA era uma estrutura bifilamentar altamente ordenada com subestruturas repetidas espaçadas a cada 0,34 nanômetro ($1 \text{ nm} = 10^{-9} \text{ metro}$) ao longo do eixo da molécula.

Com base nos dados químicos de Chargaff, nos dados de difração de Wilkins e Franklin e em deduções da construção de modelos, Watson e Crick propuseram que o DNA existe

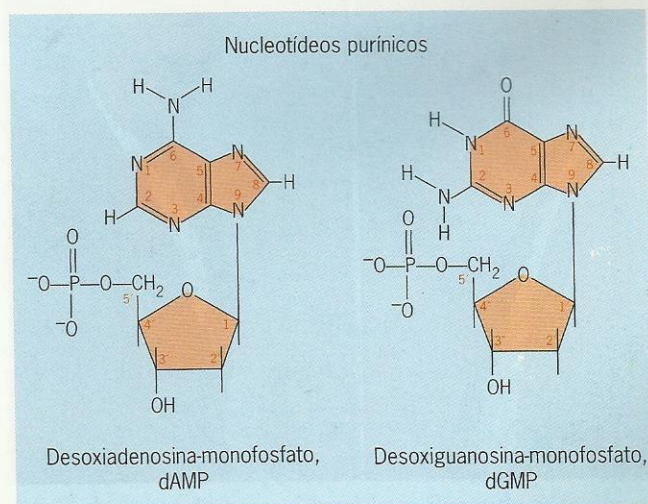
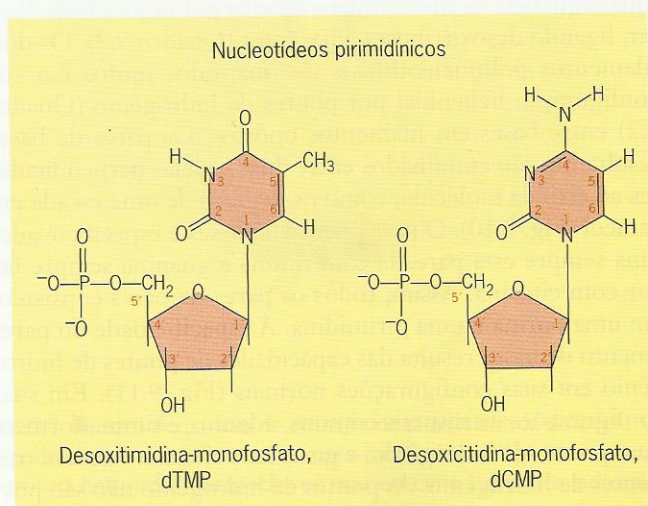


Fig. 9.6 ■ Estruturas dos quatro desoxirribonucleotídeos comuns presentes no DNA.

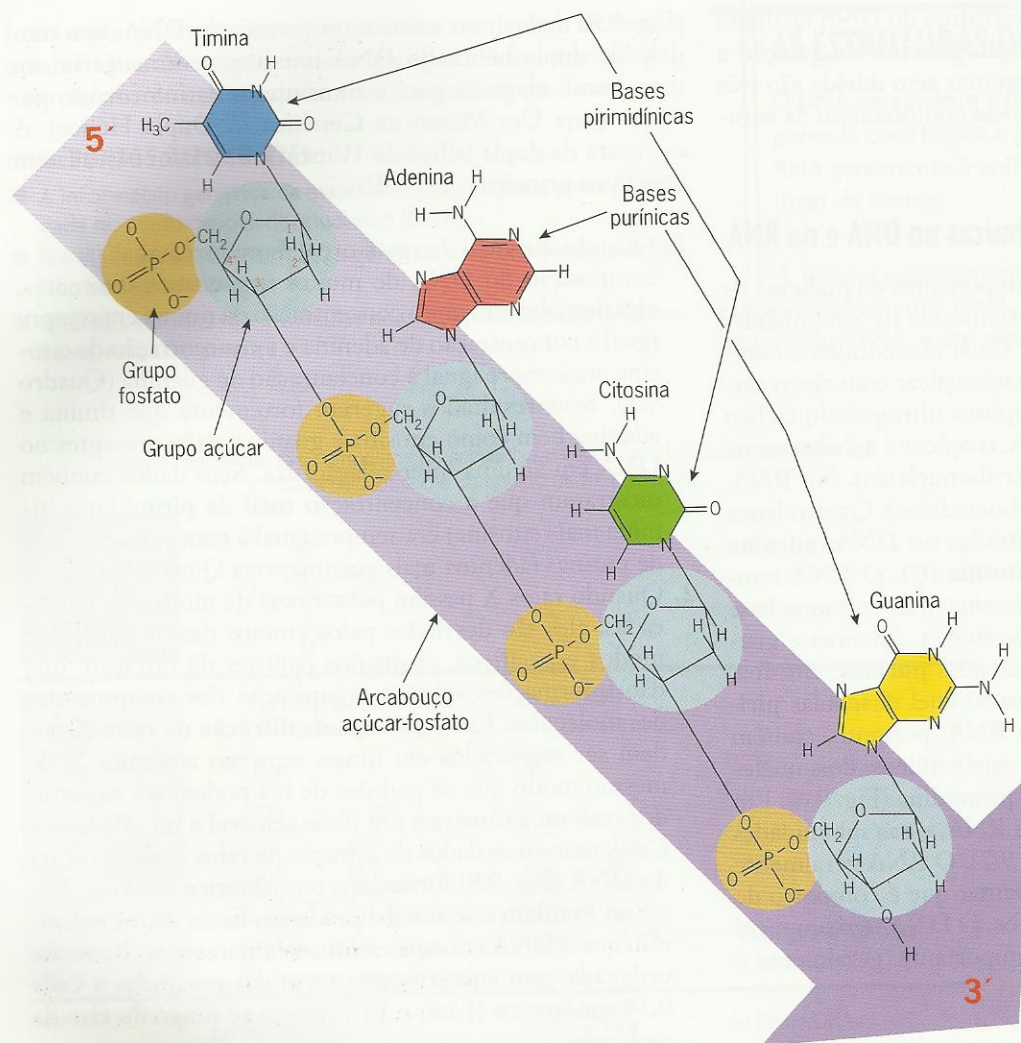


Fig. 9.7 ■ Formação de uma cadeia polinucleotídica pela união de nucleotídeos com ligações fosfodiéster. Note que o polinucleotídeo tem uma polaridade química 5' para 3' (de cima para baixo); cada ligação fosfodiéster une o carbono 5' da 2'-desoxirribose em um nucleotídeo ao carbono 3' da 2'-desoxirribose no nucleotídeo adjacente, resultando na polaridade química da cadeia polinucleotídica.

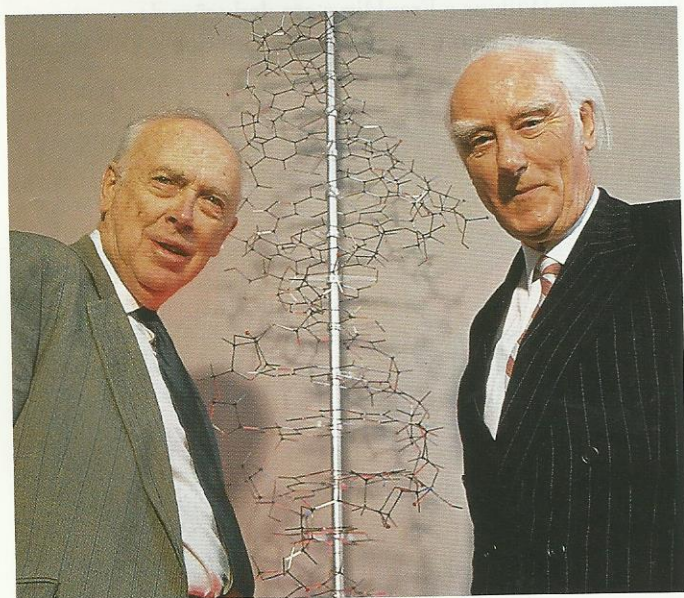


Fig. 9.8 ■ James D. Watson (esquerda) e Francis H. C. Crick com um modelo da dupla hélice de DNA.

como uma **dupla hélice** com giro para a direita em que duas cadeias polinucleotídicas são helicoidizadas uma ao redor da outra (Fig. 9.10). Cada cadeia polinucleotídica consiste em uma sequência de nucleotídeos ligados por uniões fosfodiéster, ligando desoxirriboses adjacentes (Quadro 9.2). Os dois filamentos polinucleotídicos são mantidos juntos em sua configuração helicoidal por pontes de hidrogênio (Quadro 9.2) entre bases em filamentos opostos. Os pares de bases resultantes são empilhados entre duas cadeias perpendiculares ao eixo da molécula, como os degraus de uma escada em caracol (Fig. 9.10). O pareamento de bases é específico: adenina sempre está pareada com timina e guanina sempre faz par com citosina. Assim, todos os pares de bases consistem em uma purina e uma pirimidina. A especificidade do pareamento de bases resulta das capacidades de pontes de hidrogênio em suas configurações normais (Fig. 9.11). Em suas configurações estruturais comuns, adenina e timina formam duas pontes de hidrogênio, e guanina e citosina formam três pontes de hidrogênio. As pontes de hidrogênio não são possíveis entre citosina e adenina ou timina e guanina quando elas existem em seus estados estruturais comuns.

QUADRO 9.1

Composição de Bases do DNA de Vários Organismos

Espécie	% de Adenina	% de Guanina	% de Citosina	% de Timina	Proporções Molares	
					$\frac{A + G}{T + C}$	$\frac{A + T}{G + C}$
I. Vírus						
Bacteriófago λ	26,0	23,8	24,3	25,8	0,99	1,08
Bacteriófago T2	32,6	18,1	16,6	32,6	1,03	1,88
Herpes simples	13,8	37,7	35,6	12,8	1,06	0,36
II. Bactérias						
<i>Escherichia coli</i>	26,0	24,9	25,2	23,9	1,04	1,00
<i>Micrococcus lysodeikticus</i>	14,4	37,3	34,6	13,7	1,07	0,39
<i>Ramibacterium ramosum</i>	35,1	14,9	15,2	34,8	1,00	2,32
III. Fungos						
<i>Neurospora crassa</i>	23,0	27,1	26,6	23,3	1,00	0,86
<i>Aspergillus niger</i>	25,0	25,1	25,0	24,9	1,00	1,00
<i>Saccharomyces cerevisiae</i>	31,7	18,3	17,4	32,6	1,00	1,80
IV. Eucariontes Superiores						
<i>Zea mays</i> (milho)	25,6	24,5	24,6	25,3	1,00	1,04
<i>Drosophila melanogaster</i>	30,7	19,6	20,2	29,4	1,01	1,51
<i>Homo sapiens</i> (humanos)	30,2	19,9	19,6	30,3	1,01	1,53

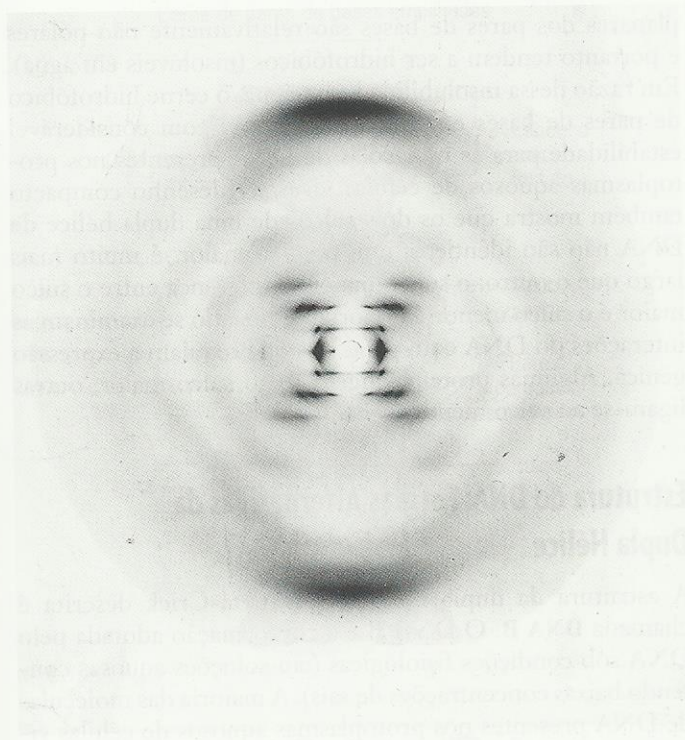


Fig. 9.9 ■ Foto do padrão de difração de raios X obtida do DNA. O padrão central em cruz indica que a molécula de DNA tem uma estrutura helicoidal e as bandas escuras em cima e embaixo indicam que as bases estão empilhadas perpendicularmente ao eixo da molécula com periodicidade de 0,34 nm.

Uma vez conhecida a seqüência de bases de um filamento da dupla hélice de DNA, a seqüência de bases do outro filamento também é conhecida devido ao pareamento específico de bases. Os dois filamentos de uma dupla hélice de DNA portanto são ditos complementares. *Esta propriedade, a complementariedade dos dois filamentos da dupla hélice, torna o DNA adequado para estocar e transmitir informação genética de geração a geração* (Cap. 10).

Os pares de bases no DNA são empilhados distando cerca de 0,34 nm, com 10 pares de bases por giro (360°) da dupla hélice (Fig. 9.10). As estruturas de açúcar-fosfato dos dois filamentos complementares são **antiparalelas** (Fig. 9.11). Unidirecionalmente ao longo da dupla hélice de DNA, as ligações fosfodiéster em um filamento vão do carbono 3' de um nucleotídeo para o carbono 5' do nucleotídeo adjacente, enquanto aquelas presentes no filamento complementar vão do carbono 5' para um carbono 3'. Essa "polaridade oposta" dos filamentos complementares de uma dupla hélice de DNA tem um papel importante em replicação, transcrição e recombinação do DNA.

A estabilidade da dupla hélice de DNA resulta em parte do maior número de pontes de hidrogênio entre os pares de bases (muito embora cada ponte de hidrogênio em si seja fraca, mais fraca que uma ligação covalente) e em parte de ligações hidrofóbicas (ou forças de empilhamento) entre pares de bases adjacentes (Quadro 9.2). A natureza do empilhamento de pares de bases é mais bem ilustrada com um diagrama compacto da estrutura do DNA (Fig. 9.12). Os lados

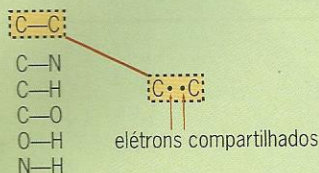
QUADRO 9.2

Ligações Químicas Importantes na Estrutura do DNA

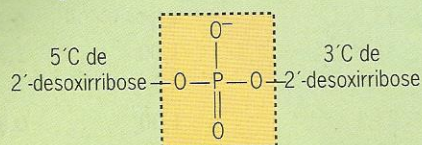
(a) Ligações covalentes

Fortes ligações químicas formadas por compartilhamento de elétrons entre átomos.

(1) Em bases e açúcares

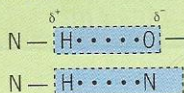


(2) Em ligações fosfodiéster



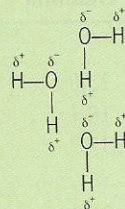
(b) Pontes de hidrogênio

Uma ligação fraca entre um átomo eletronegativo e um átomo de hidrogênio (eletropositivo) que está covalentemente ligado a um segundo átomo eletronegativo.



(c) "Ligações" hidrofóbicas

A associação de grupos não polares uns com os outros quando presentes em soluções aquosas devido a sua insolubilidade em água.



Moléculas de água são muito polares (δ^- O e δ^+ H). Compostos que são similarmente polares são muito solúveis em água ("hidrofílicos"). Compostos que são não polares (sem grupos carregados) são muito insolúveis em água ("hidrofóbicos").

Os pares de bases empilhados fornecem um cerne hidrofóbico.

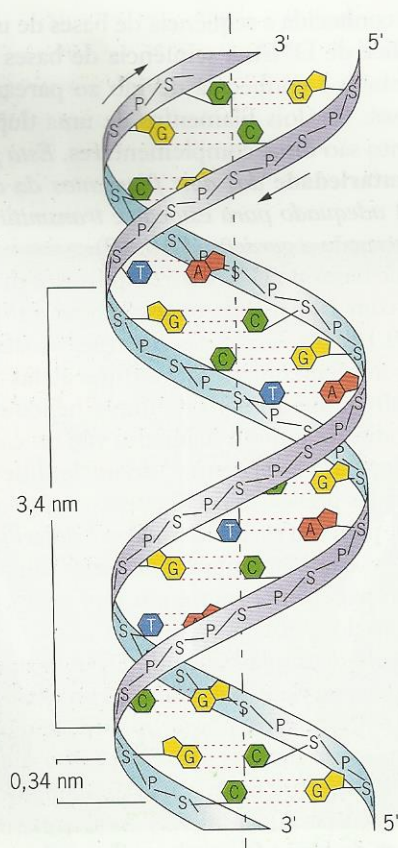
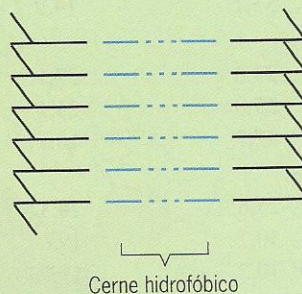


Fig. 9.10 ■ Diagrama da estrutura de dupla hélice do DNA.

planares dos pares de bases são relativamente não-polares e portanto tendem a ser hidrofóbicos (insolúveis em água). Em razão dessa insolubilidade em água, o cerne hidrofóbico de pares de bases empilhados contribui com considerável estabilidade para as moléculas de DNA presentes nos protoplasmas aquosos de células vivas. O desenho compacto também mostra que os dois sulcos de uma dupla hélice de DNA não são idênticos; um, o sulco maior, é muito mais largo que o outro, o sulco menor. A diferença entre o sulco maior e o sulco menor é importante quando se examinam as interações do DNA com proteínas que regulam a expressão gênica. Algumas proteínas ligam-se ao sulco maior; outras ligam-se ao sulco menor.

Estrutura do DNA: Formas Alternativas da Dupla Hélice

A estrutura da dupla hélice de Watson-Crick descrita é chamada **DNA B**. O DNA B é a conformação adotada pelo DNA sob condições fisiológicas (em soluções aquosas contendo baixas concentrações de sais). A maioria das moléculas de DNA presentes nos protoplasmas aquosos de células vivas existe na conformação B. Entretanto, o DNA não é uma molécula estática invariante. Pelo contrário, as moléculas de DNA exibem considerável flexibilidade conformacional.

As estruturas das moléculas de DNA alteram-se em função de seu ambiente. A conformação exata de determinada

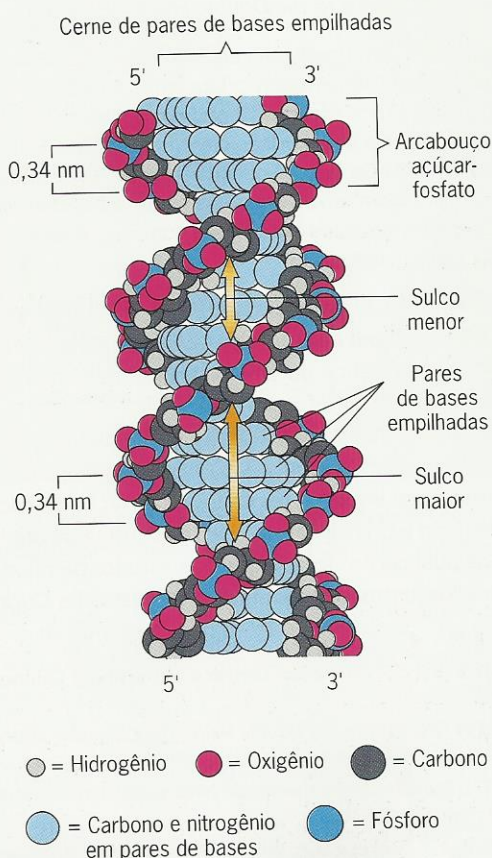
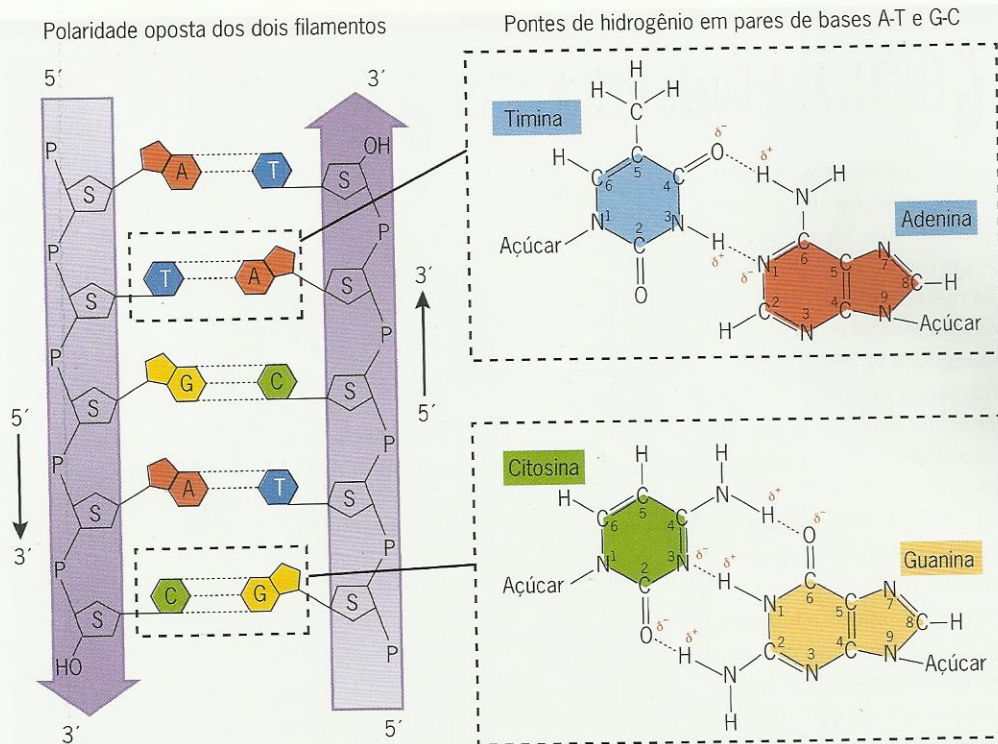


Fig. 9.12 ■ Diagrama compacto de uma dupla hélice de DNA.

molécula de DNA ou segmento de uma molécula de DNA irá depender da natureza da molécula com a qual interage. De fato, o DNA B parece ter uma média de 10,4 pares de nucleotídeos por giro, e não exatamente 10 como mostrado na Fig. 9.10. Em altas concentrações de sais ou em estado parcialmente desidratado, o DNA existe como **DNA A**, que é uma hélice com giro para a direita como o DNA B, mas com 11 pares de nucleotídeos por giro (Quadro 9.3). O DNA A é uma dupla hélice mais espessa e mais curta com diâmetro de 2,3 nm. As moléculas de DNA quase que certamente nunca existem como DNA A *in vivo*. Entretanto, a conformação DNA A é importante porque heterodúplexes DNA-RNA (hélices duplas contendo um filamento de DNA fazendo pares de bases com um filamento complementar de RNA) ou dúplexes RNA-RNA existem em uma estrutura muito similar *in vivo*.

Foi demonstrado que existem algumas seqüências de DNA com giro para a esquerda, uma forma de dupla hélice

QUADRO 9.3

Formas Alternativas de DNA

Forma da Hélice	Direção da Hélice	Pares de Bases por Giro	Diâmetro da Hélice
A	Giro para a direita	11	2,3 nm
B	Giro para a direita	10	1,9 nm
Z	Giro para a esquerda	12	1,8 nm



UM MARCO NA GENÉTICA: A Dupla Hélice

Os primeiros geneticistas disseram-nos onde estão os genes, nos cromossomos, que eles controlam os fenótipos e que eles podem passar por mutações para formas diferentes, mas não puderam dizer-nos de que são feitos os genes. Os primeiros *insights* sobre a natureza química dos genes surgiram na metade do século vinte. Em 1944, Avery, MacLeod e McCarty mostraram que o DNA é responsável por transformação em *S. pneumoniae* e, em 1952, Hershey e Chase demonstraram que a informação genética do fago T2 está presente no DNA. Assim, os genes parecem ser feitos de DNA. Entretanto, ninguém conhecia a estrutura do DNA ou como o DNA podia estocar e transmitir informação genética.

Em 25 de abril de 1953, James Watson e Francis Crick esclareceram tudo quando publicaram um trabalho descrevendo a estrutura da dupla hélice de DNA (Fig. 1) no periódico inglês *Nature*¹. Essa publicação resumida, pouco mais de uma página, mudou a biologia para sempre. Nenhuma outra publicação teve um impacto comparável nas ciências biológicas. A descoberta de Watson e Crick sobre a dupla hélice foi o começo de uma revolução na biologia molecular, a revolução que hoje nos deu as seqüências de genomas inteiros, incluindo dois rascunhos da seqüência do genoma humano.

O que há de tão marcante sobre a dupla hélice de Watson e Crick? A complementariedade dos dois filamentos – adenina em um filamento faz pontes de hidrogênio com timina no outro filamento e guanina em um filamento faz pontes de hidrogênio com citosina no outro filamento – sugeriu claramente um mecanismo pelo qual o material genético podia ser duplicado e transferido de células parentais para células filhas. Em sua publicação, Watson e Crick fizeram uma menção a esta perspectiva: “Não deixamos de notar que o pareamento específico que postulamos sugere imediatamente um possível mecanismo de cópia para o material genético.”² Cinco semanas depois, eles apresentaram seu mecanismo para a replicação do DNA em outro trabalho na *Nature*³, e em 1958 seu mecanismo de “molde” de replicação do DNA foi demonstrado como correto por Matthew Meselson e Franklin Stahl (veja Cap. 10).

Entretanto, o significado da descoberta de Watson e Crick não foi universalmente reconhecido naquela época. Alguns geneticistas pensaram que o DNA era quimicamente muito simples para explicar toda a diversidade fenotípica evidente no mundo. Eles questionaram que as proteínas, quimicamente mais complexas, tinham um melhor ajuste à variabilidade fenotípica observada. De fato, em 1955 (dois anos após a publicação do trabalho de Watson e Crick), Richard Goldschmidt, professor emérito da University of California, Berkeley, e um dos geneticistas líderes dessa geração, escreveu em apoio à idéia de que os genes são feitos de proteínas em seu livro *Theoretical Genetics*:

¹Watson, J. D., e F. H. C. Crick 1953. Molecular structure of nucleic acids. A structure for deoxyribonucleic acid. *Nature* 171:737-738.

²Ibid.

³Watson, J. D., e F. H. C. Crick. 1953. Genetical implications of the structure of deoxyribonucleic acid. *Nature* 171: 964-967.



Fig. 1 ■ O diagrama da estrutura do DNA na publicação de Watson e Crick que ganhou o Prêmio Nobel. A legenda original diz: “Esta figura é puramente diagramática. As duas fitas representam as duas cadeias fosfato-açúcar e os bastões horizontais os pares de bases unindo as cadeias. A linha vertical marca o eixo da fibra.”

Minha conclusão a partir dos fatos disponíveis hoje é de que a proteína do cromonema é o próprio material genético, mas requer moléculas ligadas de DNA para sua replicação.⁴

E Goldschmidt não foi a única pessoa cética quanto ao significado da dupla hélice. Em um suplemento especial da *Nature* publicado no jubileu do aniversário da publicação de Watson e Crick, Robert Olby, autor de *The Path to the Double Helix*⁵, escreveu:

As descobertas anteriores em geral são engrandecidas em retrospectos, especialmente em celebrações de jubileu, e a dupla hélice não é uma exceção. O registro histórico revela uma resposta muda pela comunidade científica à proposição desta estrutura em 1953. De fato, só quando surgiram os destaques de um mecanismo para envolvimento do DNA na síntese de proteínas que a comunidade bioquímica começou a ter um sério interesse por esta estrutura.⁶

Entretanto, em 1962, quando Watson, Crick e Maurice Wilkins dividiram o Prêmio Nobel de Fisiologia ou Medicina, todos reconheceram que a dupla hélice do DNA era perfeitamente adequada para estocar e transmitir informação genética.

Por que Wilkins compartilhou o Prêmio Nobel com Watson e Crick? Wilkins, Herbert Wilson e Alec Stokes do King's College em Londres forneceram importantes dados de difração de raios X que ajudaram Watson e Crick a desenvolver a estrutura da dupla hélice do DNA. Sua publicação⁷ foi um dos dois estudos de raios X sobre DNA publicados juntos com o trabalho de Watson e Crick no nú-

⁴Goldschmidt, R. B. 1955. *Theoretical Genetics*. University of California Press, Berkeley, p. 57.

⁵Olby, R. C. 1974. *The Path to the Double Helix. The Discovery of DNA*. Macmillan Press, London.

⁶Olby, R. C. 1974. Quiet debut for the double helix. p. 402. Em *The Double Helix – 50 years*. *Nature* 421: 402-405.

⁷Wilkins, M. H. F., A. R. Stokes e H. R. Wilson. 1953. Molecular structure of deoxyribose nucleic acids. *Nature* 171: 738-740.



UM MARCO NA GENÉTICA (continuação)



Fig. 2 ■ Os quatro principais participantes, Francis Crick, Maurice Wilkins, James Watson e Rosalind Franklin (sentido horário a partir da esquerda), na descoberta da estrutura da dupla hélice de DNA.

mero de 25 de abril da *Nature*. A outra publicação⁸ foi de Rosalind Franklin e Raymond Gosling, também do King's College. Franklin era uma jovem em ascensão no campo da cristalografia de raios X. Suas fotos de raios X dos cristais de DNA eram as melhores disponíveis na época. Elas evidenciaram a estrutura helicoidal do DNA e ajudaram a determinar o espaçamento entre pares de bases.

⁸Franklin, R. E. e R. G. Gosling, 1953. Molecular configuration in sodium thymonucleate. *Nature* 171: 740-741.

O relato autobiográfico de Watson sobre a descoberta da dupla hélice⁹ foi um pouco desconsiderante de Franklin. Entretanto, algumas pessoas sugeriram que Franklin esteve perto de decifrar a estrutura do DNA. Seu estudante Aaron Klug disse:

⁹Watson, J. D. 1968. *The Double Helix: A Personal Account of the Discovery of the Structure of DNA*. Atheneum, New York.



UM MARCO NA GENÉTICA (continuação)

Em seu livro, Watson escreveu que a “aceitação instantânea” por Franklin do modelo de Watson-Crick no começo causou-lhe admiração. Mas ele disse que após uma reflexão não foi tão surpreendente para ele. Não é surpreendente quando se estudam as publicações e anotações dela e se percebe o quanto ela progrediu em seu trabalho, embora de modo desconectado em épocas diferentes, sobre várias características da estrutura contida na solução correta.¹⁰

Infelizmente, Franklin morreu de câncer de ovário em 1958 aos 37 anos de idade, e uma das regras de Alfred Nobel ao estabelecer os prêmios Nobel foi que ele não poderia ser dado postumamente. Assim, Franklin não era elegível para ser premiada em 1962, muito embora suas contribuições para a dupla hélice de Watson e Crick tivessem sido iguais às de Wilkins. Entretanto, o trabalho de todos os quatro cientistas, Watson, Crick, Wilkins e Franklin (Fig. 2), contribuiu para a elucidação da estrutura do DNA.

QUESTÕES PARA DISCUSSÕES

1. O fato de que Rosalind Franklin não dividiu o Prêmio Nobel com Watson, Crick e Wilkins às vezes é chamado de injustiça. Algumas feministas consideram sua omissão o resultado de uma tendenciosidade contra mulheres cientistas em uma disciplina dominada por homens. Claramente, havia poucas mulheres biofísicas nos anos 1950, e sem dúvida havia alguma verdade

nesta afirmação. Os fatos críticos são que Franklin morreu em 1958, muito cedo em sua carreira curta mas notável, e quatro anos antes que Watson, Crick e Wilkins tivessem recebido o Prêmio Nobel, e as regras que controlam o Nobel excluem premiações póstumas. Portanto, Franklin não era elegível para dividir o Prêmio Nobel dado pela estrutura da dupla hélice de DNA. O que você acha que o comitê de seleção do Nobel teria feito se ela ainda estivesse viva em 1962?

2. Watson e Crick basearam-se muito nos padrões de difração de raios X não publicados de Wilkins, Stokes e Wilson e de Franklin e Gosling para desenvolver a estrutura da dupla hélice do DNA. Wilkins mostrou a Watson a foto mais informativa produzida por Franklin e Gosling cerca de três meses antes que seus trabalhos fossem publicados. Um mês depois, Max Perutz mostrou a Watson e Crick sua cópia do relato do Medical Research Council sobre pesquisas em andamento, incluindo o de Wilkins, Franklin e colaboradores. A descoberta da dupla hélice, portanto, documenta os efeitos positivos de compartilhar uma informação para o progresso científico. Entretanto, a livre troca de informações também dificulta atribuir o crédito apropriado aos contribuintes. Assim, as sociedades modernas têm direitos autorais, marca registrada e leis de patentes que protegem a propriedade intelectual de uma pessoa ou grupos intelectuais. Quais são os prós e contras da livre troca de dados e idéias? Quais as vantagens e desvantagens de direito autoral, marca registrada e leis de patentes? Qual o equilíbrio apropriado entre a livre troca de informações e idéias e a proteção da propriedade intelectual de um indivíduo?

¹⁰Klug, A. 1968. Rosalind Franklin and the discovery of the structure of DNA. p. 844 em *Nature* 219: 808-810, 843-844.

chamada **DNA Z** (Z para zigzag dos arcos dos açúcares-fosfato da estrutura). O DNA Z foi descoberto por análise de difração de raios X de cristais formados por oligômeros de DNA contendo pares de bases G-C e C-G alternantes. O DNA Z ocorre em hélices duplas que são ricas em G-C e contêm purinas e pirimidinas alternantes. Além de sua estrutura única helicoidal com giro para a esquerda, o DNA Z (Quadro 9.3) difere das conformações A e B por ter 12 pares de bases por giro, diâmetro de 1,8 nm e um único sulco profundo. A função do DNA Z em células vivas ainda não está clara.

Estrutura do DNA: Super-hélices Negativas *In Vivo*

Todas as moléculas funcionais de DNA presentes em células vivas apresentam um outro nível muito importante de organização – elas são super-hélices. As **super-hélices** são introduzidas em uma molécula de DNA quando um filamento ou ambos são cortados e quando os filamentos complementares de uma ponta são girados ou torcidos ao redor um do outro

com a outra ponta fixa no espaço, e portanto não podendo girar. Essa super-hélice faz com que a molécula de DNA colapse em uma estrutura muito helicoidizada similar a um fio de telefone ou elástico torcido (Fig. 9.13, parte inferior direita). As super-hélices são introduzidas e removidas de moléculas de DNA por enzimas que têm papéis essenciais na replicação do DNA (Cap. 10) e em outros processos.

A super-hélice só ocorre em moléculas de DNA com pontas fixas, pontas que não estão livres para girar. Obviamente, as pontas das moléculas de DNA circular (Fig. 9.13) presentes na maioria dos cromossomos procarióticos e em organelas de cromossomos eucarióticos tais como mitocôndrias são fixas. As grandes moléculas lineares de DNA presentes em cromossomos eucarióticos também são fixadas por sua ligação em intervalos e às pontas de componentes não-DNA de cromossomos. Tais ligações permitem que enzimas introduzam super-hélices em moléculas lineares de DNA presentes em cromossomos eucarióticos, do modo como são incorporadas às moléculas circulares de DNA presentes na maioria dos cromossomos procarióticos.

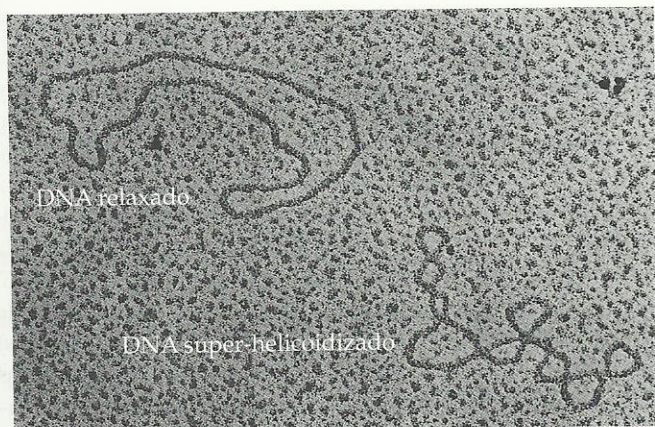


Fig. 9.13 ■ Comparação entre as estruturas do DNA relaxado e super-helicoidizado negativamente. A estrutura relaxada é o B-DNA com 10,4 pares de bases por giro da hélice. A estrutura super-helicoidizada negativamente resulta quando o B-DNA é subhelicoidizado, com menos de uma volta da hélice a cada 10,4 pares de bases.

Podemos talvez visualizar a super-hélice mais facilmente considerando uma molécula circular de DNA. Se cortarmos um filamento de uma dupla hélice circular de DNA covalentemente fechada e girarmos uma ponta do filamento cortado em um giro completo (360°) ao redor do filamento complementar enquanto seguramos a outra ponta fixada, induziremos uma super-hélice na molécula (Fig. 9.14). Se girarmos a ponta livre na mesma direção em que a dupla hélice de DNA é enrolada (giro para a direita), será produzida uma super-hélice positiva (DNA super-helicoidizado). Se girarmos a ponta livre no sentido oposto (para a esquerda), ocorrerá uma super-hélice negativa (sub-helicoidizada). Embora este seja o modo mais simples de definir super-helicoidização no DNA, não é o mecanismo pelo qual super-hélices são produzidas no DNA *in vivo*. Este mecanismo é discutido no Cap. 10.

As moléculas de DNA de quase todos os organismos, desde os menores vírus até os maiores eucariontes, exibem **super-helicoidização negativa** *in vivo*, e muitas das funções biológicas dos cromossomos podem ser feitas apenas quando

as moléculas de DNA participantes estão negativamente super-helicoidizadas. (O DNA de alguns vírus que infectam Archaea é positivamente super-helicoidizado.) Evidências consideráveis indicam que super-helicoidização negativa está envolvida em replicação (Cap. 10), recombinação, expressão gênica e regulação da expressão gênica. Quantidades similares de super-helicoidização negativa existem nas moléculas de DNA presentes em cromossomos bacterianos e cromossomos eucarióticos.

PONTOS IMPORTANTES

- O DNA geralmente existe como uma dupla hélice, sendo os dois filamentos mantidos juntos por pontes de hidrogênio entre bases complementares: adenina pareada com timina e guanina pareada com citosina.
- A complementariedade dos dois filamentos de uma dupla hélice torna o DNA adequado de modo único para estocar e transmitir informação genética.
- Os dois filamentos de uma dupla hélice de DNA têm polaridade química oposta.
- O RNA geralmente existe como uma molécula unifilar contendo uracila em vez de timina.
- As moléculas funcionais de DNA nas células são negativamente super-helicoidizadas.

ESTRUTURA CROMOSSÔMICA EM PROCARIONTES E VÍRUS

As moléculas de DNA de procariontes e vírus são organizadas em domínios de super-hélice negativa.

Grande parte da informação sobre a estrutura do DNA veio de estudos em procariontes, primariamente porque eles são menos complexos, tanto genética quanto bioquimicamente, que eucariontes. Os procariontes são monoplóides (*mono* = um); eles têm apenas um conjunto de genes (uma cópia do genoma). (“Monoplóide” não deve ser confundido com “haplóide”, que se refere especificamente ao número

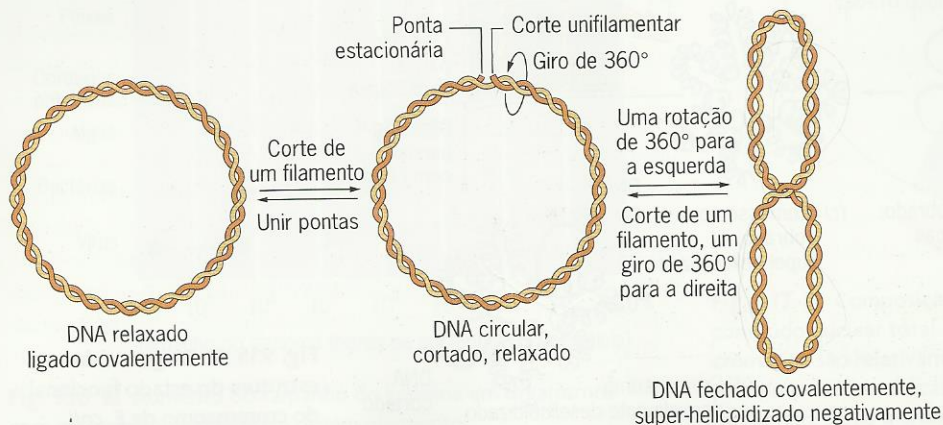


Fig. 9.14 ■ Uma definição visual do DNA super-helicoidizado negativamente. Embora a estrutura das super-hélices do DNA seja mais claramente ilustrada pelo mecanismo mostrado aqui, as super-hélices de DNA são produzidas por um mecanismo diferente *in vivo* (veja Cap. 10).

reduzido de cromossomos em gametas.) Na maioria dos vírus e procariontes, o conjunto único de genes é estocado em um só cromossomo, que por sua vez contém uma única molécula de ácido nucleico (seja RNA ou DNA).

Os menores vírus com RNA conhecidos têm apenas três genes, e as seqüências completas de nucleotídeos dos genomas de muitos vírus são conhecidas. Por exemplo, a única molécula de RNA no genoma do bacteriófago MS2 consiste em 3.569 nucleotídeos e contém 4 genes. Os menores vírus com DNA conhecidos têm apenas 9 a 11 genes. Novamente, as seqüências completas de nucleotídeos são conhecidas em vários casos. Por exemplo, o genoma do bacteriófago ϕ X174 é uma única molécula de DNA com 5.386 nucleotídeos de tamanho que contém 11 genes. Os maiores vírus com DNA, como o bacteriófago T2 e os vírus pox de animais, contêm cerca de 150 genes. Bactérias como *E. coli* têm 2.500 a 3.500 genes, a maioria dos quais estão presentes em uma única molécula de DNA.

No passado, os cromossomos procarióticos eram geralmente caracterizados como “moléculas nuas de DNA”, em contraste com cromossomos eucarióticos com suas proteínas associadas e morfologia complexa. Esta concepção errada resultou em parte porque (1) a maioria das fotos publicadas de “cromossomos” procarióticos eram micrografias eletrônicas de moléculas isoladas de DNA não metabolicamente ativas ou cromossomos funcionais e (2) a maioria das fotos publicadas de cromossomos eucarióticos eram de cromossomos meióticos ou mitóticos altamente condensados, novamente estados cromossômicos metabolicamente inativos. Cromossomos procarióticos funcionais, ou nucleóides (nucleóides, e não núcleos, porque não são delimitados por membrana nuclear), hoje comprovadamente têm pouca semelhança com as moléculas de DNA isoladas de vírus e bactérias vistas em micrografias eletrônicas, como os cromossomos interfási-

cos metabolicamente ativos de eucariontes têm pouca semelhança morfológica com cromossomos metafásicos mitóticos ou meióticos.

O tamanho da molécula de DNA circular presente no cromossomo da bactéria *Escherichia coli* é de cerca de 1.500 μ m. Como uma *E. coli* tem um diâmetro de apenas 1 a 2 μ m, a grande molécula de DNA presente em cada bactéria deve existir em uma configuração altamente condensada (dobrada ou helicoidizada). Quando os cromossomos de *E. coli* são isolados por procedimentos suaves na ausência de detergentes iônicos (comumente usados para lisar células) e são mantidos na presença de alta concentração de cátions tais como poliaminas (pequenas proteínas básicas ou positivamente carregadas) ou sal 1M para neutralizar a carga negativa de grupos fosfato do DNA, os cromossomos permanecem em um estado altamente condensado comparável em tamanho ao nucleóide *in vivo*. Esta estrutura, chamada **genoma compactado**, é o estado funcional de um cromossomo bacteriano. Embora menores, os cromossomos funcionais intracelulares de vírus bacterianos são muito similares aos genomas compactados de bactérias.

Dentro do genoma compactado, a grande molécula de DNA em um cromossomo de *E. coli* é organizada em 50 a 100 **domínios** ou alças, cada um dos quais é super-helicoidizado negativamente de modo independente (Fig. 9.15). RNA e proteínas são ambos componentes do genoma compactado, que podem ser parcialmente relaxados por tratamento com desoxirribonuclease (DNase) ou ribonuclease (RNase). Como cada domínio do cromossomo é independentemente super-helicoidizado, a introdução de “cortes” unifilamentares no DNA por tratamento dos cromossomos com uma DNase que corta o DNA em sítios internos irá relaxar o DNA apenas nos domínios cortados, e todas as alças não cortadas permanecerão super-helicoidizadas. A destrui-

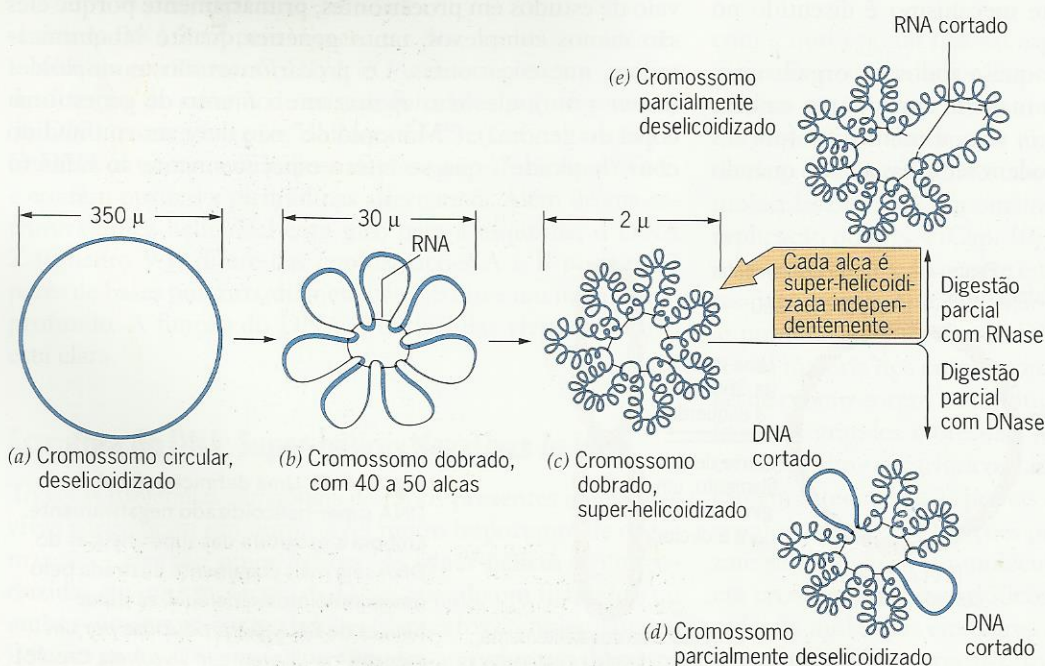


Fig. 9.15 ■ Diagrama da estrutura do estado funcional do cromossomo de *E. coli*.

ção dos conectores de RNA pela RNase descompactará o genoma compactado parcialmente eliminando a organização da molécula de DNA em 50 a 100 alças. Entretanto, o tratamento com RNase não afetará a super-helicoidização dos domínios do cromossomo.

PONTOS IMPORTANTES

- As moléculas de DNA nos cromossomos procarióticos e virais são organizadas em domínios de super-hélice negativa.
- Os cromossomos bacterianos contêm moléculas circulares de DNA segregado em cerca de 50 domínios.

ESTRUTURA CROMOSSÔMICA EM EUKARIOTES

Cromossomos eucarióticos contêm moléculas enormes de DNA que estão altamente condensadas durante mitose e meiose. Os centrômeros e telômeros dos cromossomos eucarióticos têm estruturas únicas.

Os genomas eucarióticos contêm níveis de complexidade que não são encontrados em procariotes. Em contraste com procariotes, a maioria dos eucariotes são diplóides, tendo dois conjuntos completos de genes, um de cada genitor. Como foi discutido no Cap. 6, algumas plantas com flores são poliplóides, isto é, possuem várias cópias do genoma. Embora eucariotes tenham apenas 2 a 15 vezes mais genes que a *E. coli*, eles têm ordens de grandeza maior que a do DNA (Fig. 9.16). Além disso, grande parte desse DNA não contém genes, pelo menos não genes que codificam proteínas ou moléculas de RNA.

Não só a maioria dos eucariotes possui muitas vezes a quantidade de DNA encontrada nos procariotes, mas também esse DNA está compactado em vários cromossomos, e

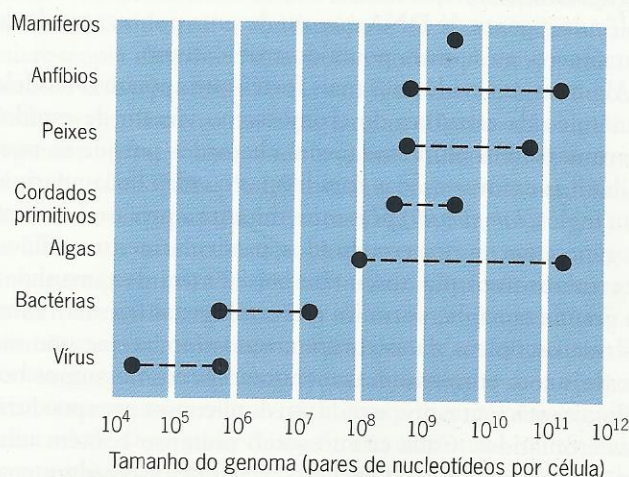


Fig. 9.16 ■ Aumento do tamanho do genoma em organismos com maior complexidade do desenvolvimento.

cada cromossomo está presente em duas (diplóides) ou mais (poliplóides) cópias. Lembre que o cromossomo de *E. coli* tem tamanho de 1.500 μm , ou cerca de 1,5 mm. Considere agora que o complemento cromossômico haplóide, ou genoma, de um humano contém cerca de 1.000 mm de DNA (ou cerca de 2.000 mm por célula diplóide). Além disso, esta medida de DNA é subdividida entre os 23 cromossomos de tamanhos e formas variadas, cada cromossomo contendo de 15 a 85 mm de DNA. No passado, os geneticistas tinham pouca informação sobre como esse DNA estava disposto nos cromossomos. Há uma molécula de DNA por cromossomo como nos procariotes ou existem muitas? Se forem muitas, como as moléculas estão dispostas em relação umas às outras? Como os 85 mm (85.000 μm) de DNA no maior cromossomo humano ficam condensados em uma estrutura metafásica mitótica que tem cerca de 0,5 μm de diâmetro e 10 μm de tamanho? Quais são as estruturas dos cromossomos interfásicos metabolicamente ativos? Consideraremos as respostas a algumas destas perguntas nas seções seguintes.

Composição Química de Cromossomos Eucarióticos

Os cromossomos interfásicos geralmente não são visíveis ao microscópio óptico. Entretanto, análise química, microscopia eletrônica e estudos de difração de raios X de **cromatina** isolada (o complexo do DNA, proteínas cromossômicas e outros constituintes cromossômicos isolados dos núcleos) deram valiosas informações sobre a estrutura de cromossomos eucarióticos.

Quando a cromatina é isolada de núcleos interfásicos, os cromossomos individuais não são reconhecíveis. Em vez

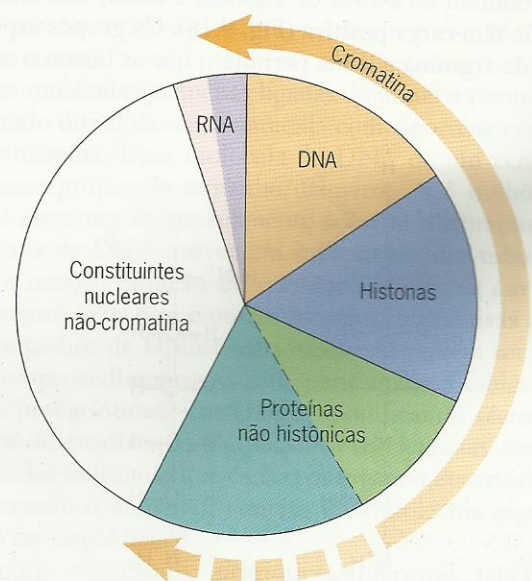


Fig. 9.17 ■ Composição química da cromatina em função do conteúdo nuclear total. Os conteúdos de DNA e histona da cromatina são relativamente constantes, mas a quantidade de proteínas não histônicas presentes depende do procedimento usado para isolar a cromatina (seta tracejada).

disso, observa-se um agregado irregular de nucleoproteínas. A análise química de cromatina isolada mostra que ela consiste primariamente em DNA e proteínas com menores quantidades de RNA (Fig. 9.17). As proteínas são de duas classes principais: (1) proteínas básicas (de carga positiva em pH neutro) chamadas **histonas** e (2) um grupo heterogêneo de proteínas ácidas (carga negativa em pH neutro) coletivamente chamadas **proteínas cromossômicas não-histonas**.

As histonas têm um importante papel estrutural na cromatina. Elas estão presentes na cromatina de todos os eucariontes em quantidades equivalentes às quantidades de DNA. Esta relação sugere que ocorre uma interação entre histonas e DNA que é conservada em eucariontes. As histonas de todas as plantas e animais consistem em cinco classes de proteínas. Esses cinco tipos principais de histonas, chamadas **H1**, **H2a**, **H2b**, **H3** e **H4**, estão presentes em quase todos os tipos de células. Existem algumas exceções, mais notadamente alguns espermatozoides, em que as histonas são substituídas por outra classe de pequenas proteínas básicas chamadas **protaminas**.

Os cinco tipos de histonas estão presentes em proporções molares de aproximadamente 1 H1:2 H2a:2 H2b:2 H3:2 H4. Quatro dos cinco tipos de histonas estão especificamente associadas ao DNA para produzir as subunidades estruturais básicas de cromatina, pequenas contas elipsóides (aproximadamente 11 nm de diâmetro por 6 nm de altura) chamadas **nucleossomos**. As histonas têm sido altamente conservadas durante a evolução – quatro dos cinco tipos de histonas são similares em todos os eucariontes.

A maioria dos 20 aminoácidos nas proteínas são de carga neutra, isto é, não têm carga em pH 7. Entretanto, algumas são básicas e algumas são ácidas. As histonas são básicas porque contêm 20 a 30% de arginina e lisina, dois aminoácidos que têm carga positiva (Fig. 9.18). Os grupos expostos $-NH_3^+$ de arginina e lisina permitem que as histonas atuem

como polications. Os grupos laterais de carga positiva nas histonas são importantes em sua interação com DNA, que é polianiónico devido aos grupos fosfato de carga negativa.

A notável constância das histonas H2a, H2b, H3 e H4 em todos os tipos de células de um organismo e mesmo entre espécies amplamente divergentes é consistente com a idéia de que são importantes na estrutura da cromatina (acondicionamento do DNA) e estão apenas inespecificamente envolvidas na regulação da expressão gênica. Entretanto, como será discutido mais adiante, as modificações químicas das histonas podem alterar a estrutura cromossômica, a qual, por sua vez, pode acentuar ou diminuir a expressão de genes situados na cromatina modificada.

Em contraste, a fração protéica não histônica da cromatina consiste em um grande número de proteínas heterogêneas. Além disso, a composição da fração de proteínas cromossômicas não histônicas varia amplamente entre diferentes tipos de células do mesmo organismo. Assim, as proteínas cromossômicas não histônicas provavelmente não têm papéis centrais na compactação do DNA nos cromossomos. Em vez disso, são prováveis candidatas a papéis na regulação da expressão de genes ou grupos de genes específicos.

Uma Grande Molécula de DNA por Cromossomo

Um cromossomo eucariótico típico contém de 1 a 20 cm (10^4 a 2×10^5 μm) de DNA. Durante a metáfase da meiose e da mitose, esse DNA está compactado em um cromossomo com um tamanho de apenas 1 a 10 μm . Como todo esse DNA é condensado em cromossomos compactos que estão presentes durante a mitose e a meiose? Muitas moléculas de DNA correm em paralelo pelo cromossomo, o modelo “multifilamentar” ou **multinêmico**, ou há apenas uma dupla hélice de DNA de uma ponta do cromossomo até a outra, o modelo unifilamentar ou **uninêmico**? (Note que “filamento” aqui se refere à dupla hélice de DNA, e não a cadeias polinucleotídicas individuais de DNA.) Evidências consideráveis agora indicam que cada cromossomo contém uma única molécula gigante de DNA que vai de uma ponta, através do centrômero, até a outra ponta do cromossomo.

Algumas das evidências mais fortes que apoiam o modelo uninêmico da estrutura do cromossomo vieram de estudos de **cromossomos plumosos** (assim chamados porque se assemelham às escovas usadas para limpar o vidro de lamparinas – em inglês, *lampbrush*) presentes durante a prófase I da ovogênese em muitos vertebrados, particularmente anfíbios. Os cromossomos plumosos têm até 800 μm de tamanho, e seu grande tamanho permitiu que citologistas fizessem estudos microscópicos da estrutura cromossômica que não são possíveis com cromossomos menores. Os cromossomos homólogos estão pareados, e cada um duplicou-se para produzir duas cromátides. Cada cromossomo plumoso contém uma região axial central, onde as duas cromátides estão altamente condensadas, e vários pares de alças laterais (Fig. 9.19). As alças são regiões transcricionalmente ativas de cromátides

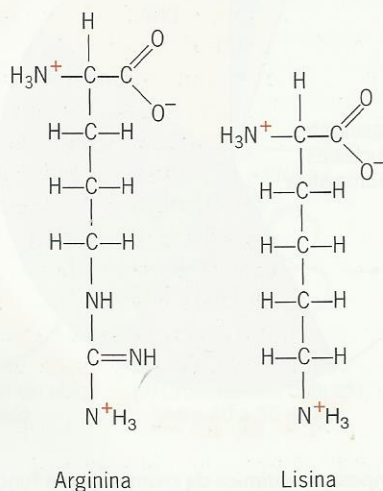


Fig. 9.18 ■ Estruturas dos aminoácidos arginina e lisina (em pH 7), que juntos correspondem a 20 a 30 por cento dos aminoácidos em histonas.

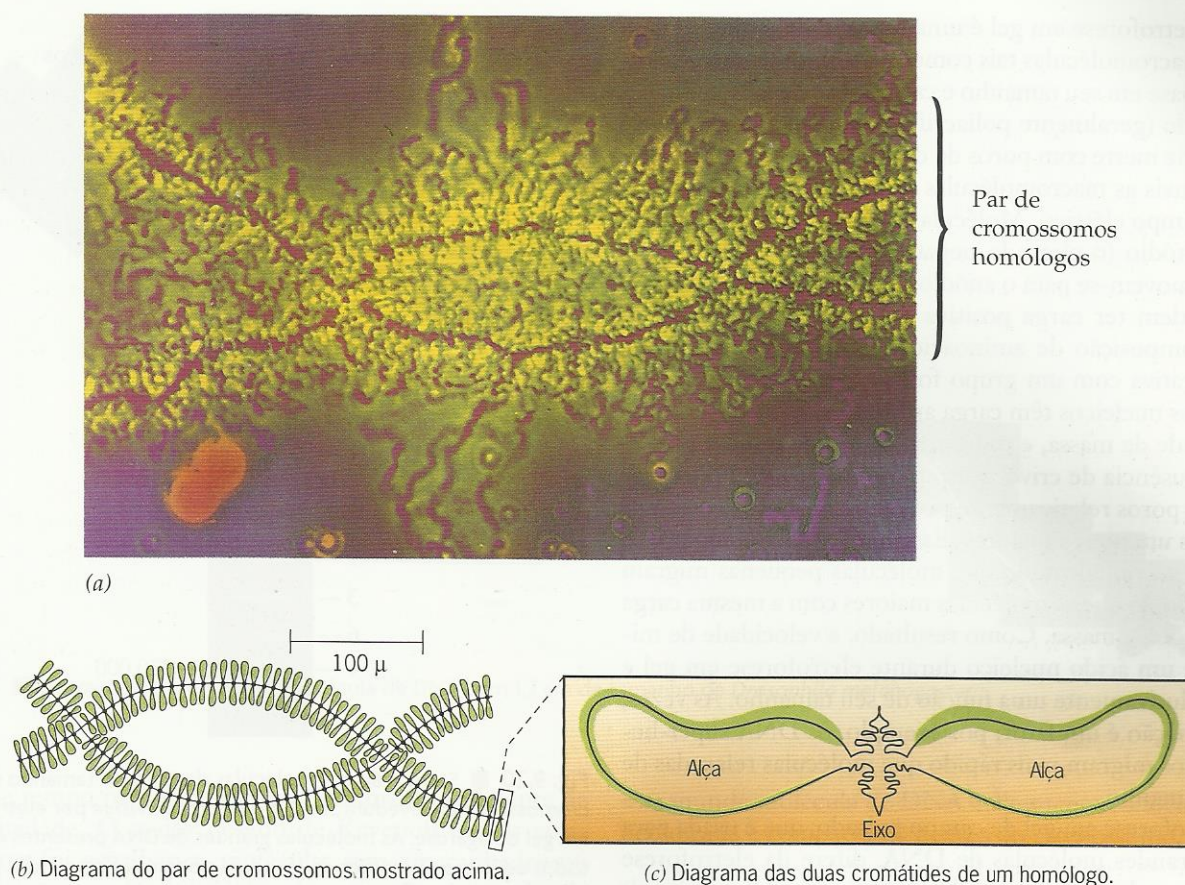


Fig. 9.19 ■ Micrografia de contraste de fase (a) e diagrama (b) de um par de cromossomos plumosos em um ovócito de tritão, *Triturus viridescens*. As estruturas das regiões axial e de alça de duas cromátides de um único cromossomo plumoso são mostradas em (c). O elemento central de cada cromátide (tanto regiões axiais quanto alças laterais) é uma única molécula de DNA. O material que circunda a molécula de DNA é o RNA nascente primário que é sintetizado no DNA distendido nas regiões de alça.

únicas. A integridade tanto do eixo central quanto das alças laterais depende do DNA. O tratamento com DNase produz quebras tanto no eixo quanto nas alças. O tratamento com RNase ou proteases remove material circundante da matriz mas não destrói a continuidade do eixo ou das alças. A microscopia eletrônica dos cromossomos plumosos tratados com RNase e protease revela um filamento central de cerca de 2 nm de diâmetro nas alças laterais. Como cada alça é um segmento de uma cromátide e como o diâmetro de uma dupla hélice de DNA é de 1,9 nm, tais cromossomos plumosos devem ser estruturas uninêmicas (Fig. 9.19).

Como os cromossomos plumosos são cromossomos da linhagem germinativa, sua estrutura é particularmente relevante para uma compreensão de fenômenos genéticos. Os cromossomos de células somáticas podem ter estruturas diferentes. Embora a maioria seja uninêmica, sabe-se que alguns, tais como os cromossomos politênicos gigantes nas glândulas salivares de *Drosophila* (Cap. 6), são estruturas multinêmicas compostas de muitas moléculas idênticas de DNA.

A dúvida sobre se os cromossomos uninêmicos dos eucariontes contêm uma única molécula grande de DNA ou

muitas moléculas pequenas ligadas de ponta a ponta tem-se mostrado difícil de ser esclarecida com rigorosas evidências experimentais. Uma molécula de DNA de um centímetro tem uma proporção tamanho/diâmetro de 5 milhões para 1. Tal estrutura é extremamente sensível a agitação. Se tal molécula de DNA estiver em solução em um tubo de ensaio, a menor vibração irá quebrar a molécula em muitos fragmentos. Por este e outros motivos, estimativas precisas dos tamanhos de DNA eucarióticos não podem ser obtidas com os procedimentos usados para analisar moléculas de DNA procarióticas. Entretanto, modificando técnicas antigas e desenvolvendo algumas novas, cientistas obtiveram evidências indicando que cada cromossomo eucariótico, não importando o tamanho, contém uma molécula gigante de DNA em dupla hélice.

Alguns organismos eucarióticos inferiores, tais como o fungo *Neurospora crassa* e a levedura *Saccharomyces cerevisiae*, têm cromossomos relativamente pequenos. Com estes organismos, foi usado um procedimento chamado **eletroforese em gel de campo pulsado** para demonstrar que cada cromossomo contém uma única molécula de DNA. A téc-

nica de eletroforese em gel é uma ferramenta poderosa para separar macromoléculas tais como proteínas e ácidos nucleicos com base em seu tamanho e carga (veja Cap. 15). Um gel semi-sólido (geralmente poliacrilamida ou agarose) fornece uma matriz inerte com poros de determinado tamanho através dos quais as macromoléculas migram quando colocadas em um campo elétrico. Moléculas de carga positiva migram para o cátodo (o eletrodo negativo) e moléculas de carga negativa movem-se para o ânodo (o eletrodo positivo). Proteínas podem ter carga positiva ou negativa, dependendo de sua composição de aminoácidos. Ácidos nucleicos têm carga negativa com um grupo fosfato por nucleotídeo. Assim, ácidos nucleicos têm carga aproximadamente constante por unidade de massa, e todos migrariam na mesma velocidade na ausência de crivo. Entretanto, os géis de poliacrilamida têm poros relativamente pequenos e os géis de agarose têm poros um pouco maiores. Esses géis atuam como crivos moleculares, de tal modo que moléculas pequenas migram mais rápido do que as moléculas maiores com a mesma carga por unidade de massa. Como resultado, a velocidade de migração de um ácido nucleico durante eletroforese em gel é quase exclusivamente uma função de seu tamanho. Às vezes, a conformação é um fator; por exemplo, os DNA super-helicoidizados migram mais rápido que moléculas relaxadas de mesmo tamanho.

A eletroforese em gel de campo pulsado, que é usada para separar grandes moléculas de DNA, difere da eletroforese padrão em gel pois, em vez de um único (unidimensional) e constante campo elétrico, dois campos elétricos separados por 90° são aplicados através do gel de modo pulsado ou alternado. Na eletroforese padrão em gel, as moléculas de DNA passam pelo gel de modo ondulado. Na eletroforese em gel de campo pulsado, a aplicação de campos elétricos intermitentes e alternados requer que as moléculas se reorientem antes de continuar a migrar pelo gel. Moléculas maiores levam mais tempo para sofrer essa reorientação e se movem mais lentamente. Como resultado, a eletroforese em gel de campo pulsado produz melhor separação de moléculas de DNA muito grandes. Quando esta técnica foi usada para separar moléculas intactas de DNA dos fungos *N. crassa* e *S. cerevisiae*, os resultados mostraram que o número de moléculas de DNA de tamanhos diferentes era igual ao número de cromossomos não homólogos nestas espécies (Fig. 9.20).

Infelizmente, as moléculas de DNA muito grandes presentes nos cromossomos de eucariontes superiores, tais como *Drosophila* e humanos não podem ser separadas mesmo em eletroforese em gel de campo pulsado. Pesquisadores usaram enfoques adicionais na tentativa de demonstrar que os cromossomos grandes de animais superiores e plantas contêm cada um uma molécula de DNA. Auto-radiografia e viscoelastometria são dois enfoques que produziram resultados importantes.

Auto-radiografia é um método para detecção e localização de isótopos radioativos em preparações citológicas ou macromoléculas por exposição a uma emulsão fotográfica que é sensível a radiação de baixa energia. A emulsão contém

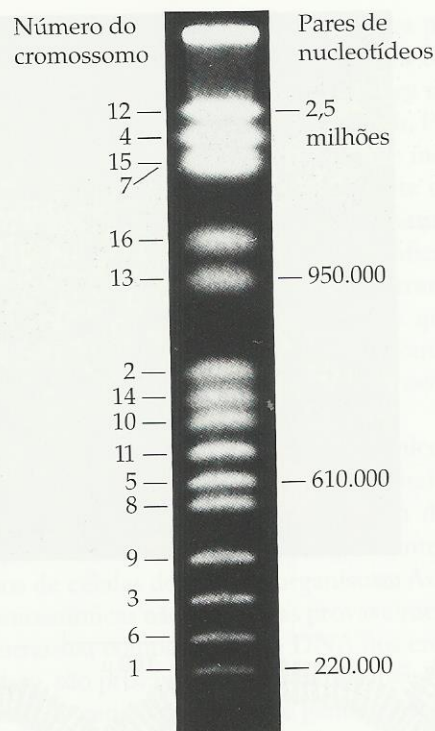


Fig. 9.20 ■ Separação das moléculas de DNA do tamanho do cromossomo da levedura *Saccharomyces cerevisiae* por eletroforese em gel de agarose. As moléculas grandes de DNA presentes em todos os 16 cromossomos da levedura podem ser resolvidas por este procedimento.

cristais de prata que produzem pequenos pontos pretos, em geral chamados grãos de prata, quando são expostos a partículas carregadas emitidas durante o decaimento de isótopos radioativos. A auto-radiografia permite que um pesquisador prepare uma imagem da localização de radioatividade em macromoléculas, células ou tecidos, como uma fotografia nos permite fazer uma imagem do que vemos. A diferença é que o filme usado para auto-radiografia é sensível a radioatividade, enquanto o filme que usamos em uma câmera é sensível a luz visível. A auto-radiografia é particularmente útil no estudo do metabolismo do DNA, pois o DNA pode ser marcado especificamente por células cultivadas em ^3H -timidina, um desoxirribonucleosídeo de timina que contém um isótopo radioativo de hidrogênio (trítio). A timidina é incorporada quase que exclusivamente ao DNA. Ela não está presente em nenhum outro componente principal da célula.

No início dos anos 1960, Ruth Kavenoff, Lynn Klotz e Bruno Zimm cultivaram células de drosófilas em um meio de cultura contendo ^3H -timidina por 24 horas, lisaram as células suavemente para não quebrar as moléculas de DNA cromossômico e cuidadosamente coletaram as moléculas de DNA em lâminas de vidro revestidas com proteínas. Eles então cobriram as lâminas com uma emulsão sensível a partículas β (os elétrons de baixa energia emitidos durante o decaimento do trítio) e as guardaram no escuro por um pe-

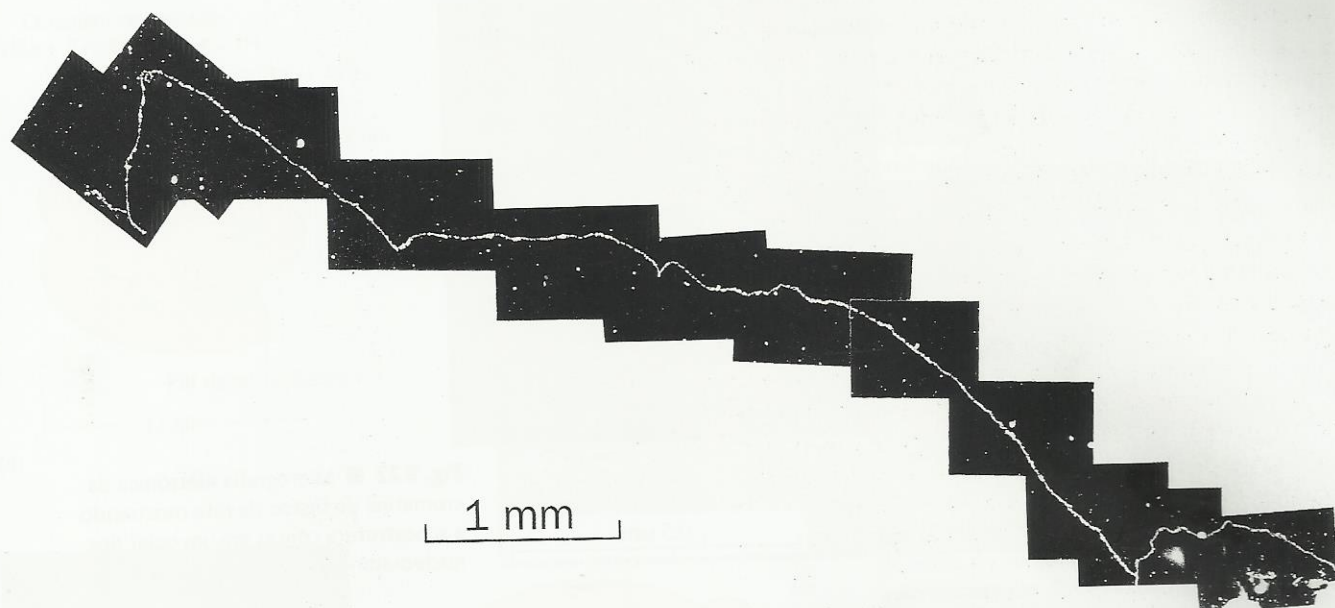


Fig. 9.21 ■ Auto-radiografia de uma molécula de DNA com 1,2 cm de tamanho de *Drosophila melanogaster*.

ríodo de tempo para permitir um decaimento suficiente de radioatividade. O maior desafio de Kavenoff e seus colaboradores foi espalhar as moléculas sem emaranhados ou superposições nas lâminas de modo que todo o tamanho de uma molécula fosse visível. Suas melhores auto-radiografias mostraram moléculas de DNA com tamanhos de até 1,2 cm (Fig. 9.21). As moléculas de DNA com este tamanho teriam uma massa de cerca de 3×10^{10} daltons (um dalton é a massa de um átomo de hidrogênio) e conteriam cerca de dois terços do DNA que se sabia estar presentes nos maiores cromossomos de *D. melanogaster*. Assim, estes resultados apoiaram o conceito das moléculas de DNA do tamanho do cromossomo em drosófilas.

Kavenoff e colaboradores também usaram uma técnica chamada **viscoelastometria**, um procedimento para analisar a viscosidade de moléculas em solução, para determinar os tamanhos das moléculas de DNA nos maiores cromossomos de drosófilas. Os dados de viscoelastometria de Kavenoff e colaboradores indicam que as maiores moléculas de DNA em drosófilas têm uma massa de $4,1 \times 10^{10}$ daltons. Como foi demonstrado que o maior cromossomo de drosófilas contém cerca de $4,3 \times 10^{10}$ daltons de DNA (total, seja uma molécula ou várias) por análise bioquímica direta, a estimativa viscoelastométrica do tamanho das maiores moléculas de DNA nos núcleos de drosófilas correlaciona-se quase exatamente com a quantidade total de DNA presente no maior cromossomo.

Este e outros resultados forneceram evidências fortes, mas não definitivas, de que cada cromossomo eucariótico contém uma longa dupla hélice de DNA que vai de uma ponta do cromossomo, através do centrômero, até a outra ponta. Entretanto, como discutiremos na seção seguinte,

essa grande molécula de DNA é altamente condensada (helicoidizada e compactada) dentro do cromossomo.

Três Níveis de Compactação do DNA em Cromossomos Eucarióticos

O maior cromossomo no genoma humano contém cerca de 85 mm ($85.000 \mu\text{m}$, ou $8,5 \times 10^7 \text{ nm}$) de DNA que se acredita existir como uma molécula gigante. Essa molécula de DNA de algum modo fica compactada em uma estrutura metafásica que tem cerca de $0,5 \mu\text{m}$ de diâmetro e cerca de $10 \mu\text{m}$ de tamanho, uma condensação de quase 10^4 vezes em tamanho desde a molécula nua de DNA até o cromossomo metafásico. Como ocorre esta condensação? Que componentes dos cromossomos estão envolvidos nos processos de compactação? Existe um esquema universal de compactação? Existem níveis diferentes de compactação? Claramente, cromossomos meióticos e mitóticos são mais extensamente condensados que os cromossomos interfásicos. Que níveis adicionais de condensação ocorrem nestas estruturas especiais que são destinadas a assegurar a segregação mais apropriada do material genético durante as divisões celulares? As seqüências de DNA dos genes que estão sendo expressas são compactadas diferentemente das dos genes que não estão sendo expressas? Vamos investigar algumas das evidências que estabeleceram a existência de três níveis diferentes de compactação do DNA em cromossomos.

Quando a cromatina isolada é examinada por microscopia eletrônica, vemos que ela consiste em uma série de contas elípticas (cerca de 11 nm de diâmetro e 6 nm de altura) unidas por finos filamentos (Fig. 9.22). Mais evi-

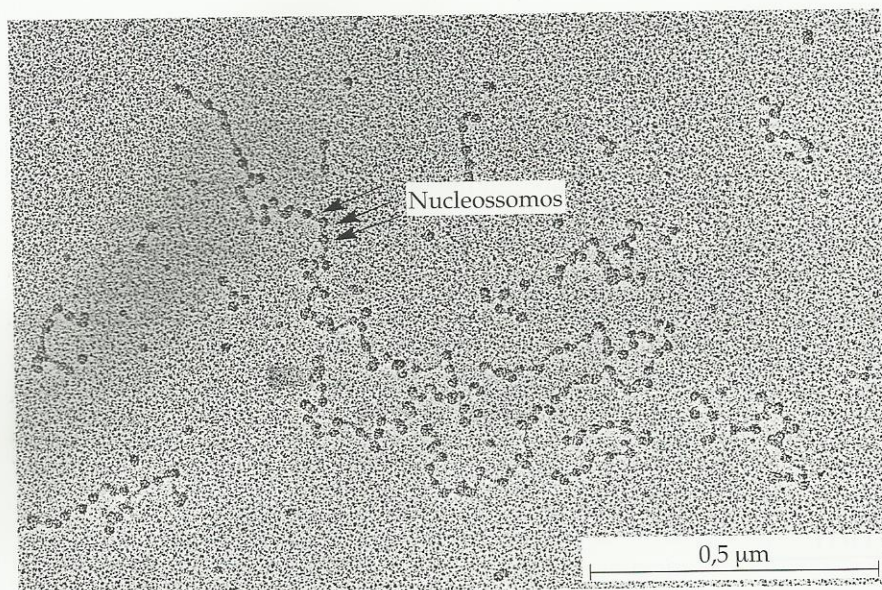


Fig. 9.22 ■ Micrografia eletrônica da cromatina de fígado de rato mostrando a subestrutura contas em um colar dos nucleossomos.

dências de uma compactação periódica regular do DNA vieram de estudos sobre a digestão da cromatina com várias nucleases. A digestão parcial da cromatina com tais nucleases produziu fragmentos de DNA em um conjunto de pequenos tamanhos que eram múltiplos integrais do fragmento de menor tamanho. Estes resultados são perfeitamente explicados se a cromatina tiver uma estrutura repetida, supostamente as contas vistas pela microscopia eletrônica (Fig. 9.22), dentro das quais o DNA é compactado sob uma forma resistente a nuclease (Fig. 9.23). Essa “conta” ou subunidade de cromatina é chamada **nucleossomo**. De acordo com o conceito atual de estrutura da cromatina, os **ligadores**, ou filamentos de DNA entre contas, são suscetíveis a ataque de nuclease.

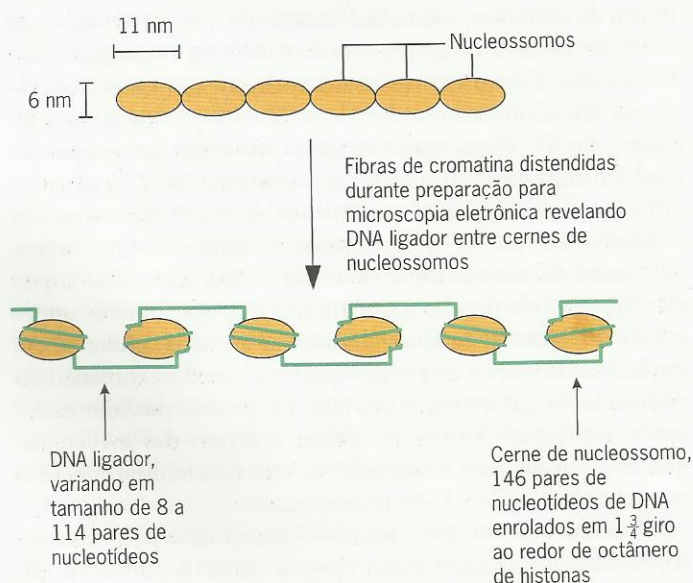


Fig. 9.23 ■ Diagrama da estrutura de cromatina do nucleossomo.

Após digestão parcial do DNA na cromatina com uma endonuclease (uma enzima que corta internamente o DNA), o DNA com aproximadamente 200 pares de nucleotídeos de tamanho está associado a cada nucleossomo (produzido por uma clivagem em cada região ligadora). Após extensa digestão por nuclease, um segmento de DNA com 146 pares de nucleotídeos de tamanho continua presente em cada nucleossomo. Esta estrutura resistente a nuclease é chamada **cerne do nucleossomo**. Sua estrutura, essencialmente invariante em eucariontes, consiste em um DNA com 146 pares de nucleotídeos de tamanho e duas moléculas cada das histonas, H2a, H2b, H3 e H4. As histonas protegem o segmento de DNA no cerne do nucleossomo de clivagem por endonucleases. Estudos físicos (difração de raios X e análises similares) de cristais do cerne do nucleossomo mostram que o DNA é enrolado com $1\frac{3}{4}$ giro de uma super-hélice ao redor do exterior do octâmero de histonas (Fig. 9.24a).

A estrutura do cerne do nucleossomo foi determinada com a resolução de 0,28 nm por estudos de difração de raios X. O mapa de alta resolução resultante do cerne do nucleossomo mostra a localização precisa de todas as oito moléculas de histona e os 146 pares de nucleotídeos da super-helicoidização negativa do DNA (Fig. 9.24b e c). Alguns dos segmentos terminais das histonas passam sobre e entre os giros da super-hélice de DNA para dar estabilidade ao nucleossomo. As interações envolvendo várias moléculas de histonas e entre as histonas e o DNA são vistas mais claramente nas estruturas de metade do cerne do nucleossomo (Fig. 9.24d), que contém apenas 73 pares de nucleotídeos de DNA super-helicoidizado.

A subunidade completa de cromatina consiste no cerne do nucleossomo, no DNA ligador e nas proteínas cromossômicas não histônicas associadas, todos estabilizados pela ligação de uma molécula de histona H1 com o exterior da estrutura (Fig. 9.24e). O tamanho do DNA ligador varia de

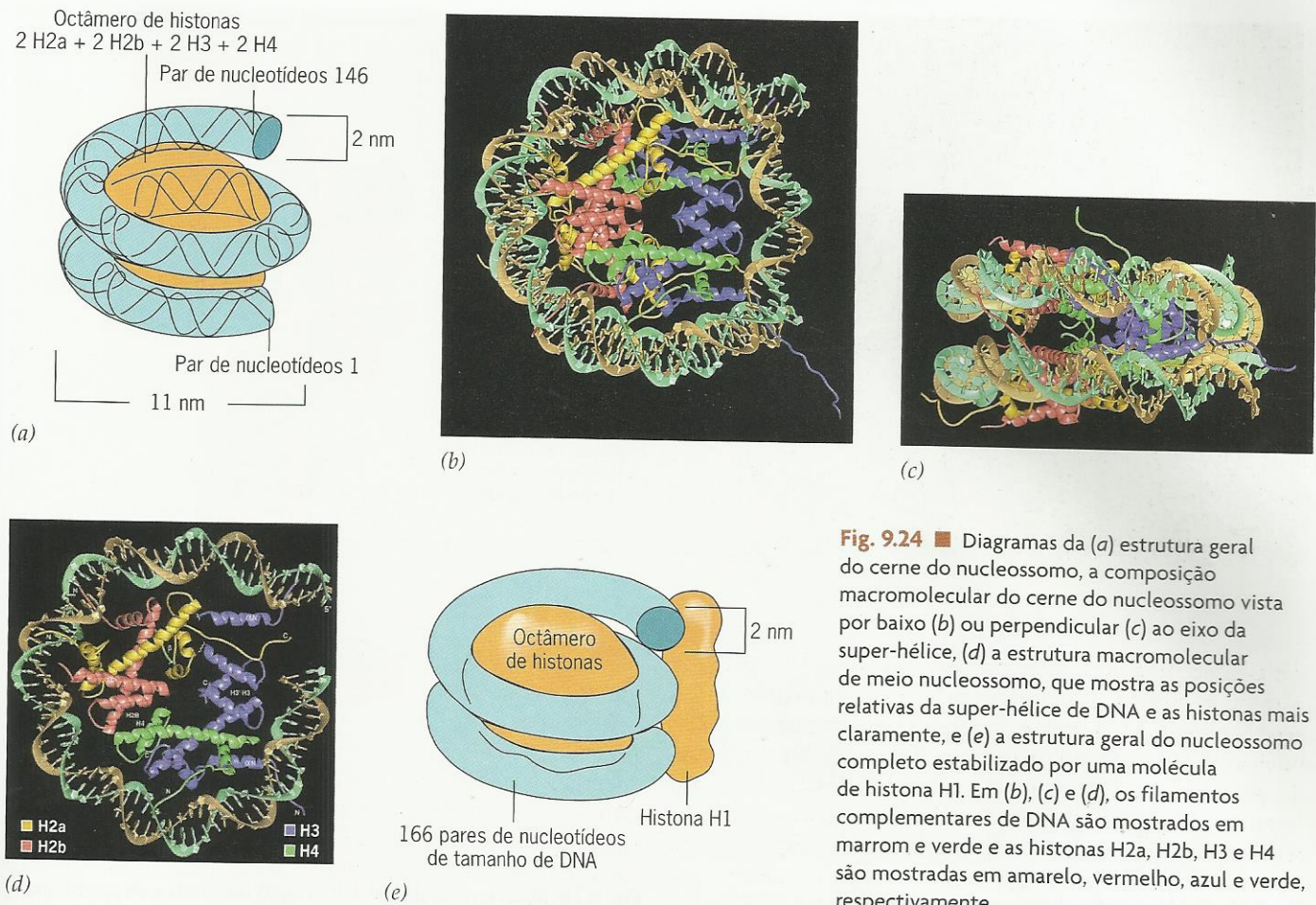


Fig. 9.24 ■ Diagramas da (a) estrutura geral do cerne do nucleossomo, a composição macromolecular do cerne do nucleossomo vista por baixo (b) ou perpendicular (c) ao eixo da super-hélice, (d) a estrutura macromolecular de meio nucleossomo, que mostra as posições relativas da super-hélice de DNA e as histonas mais claramente, e (e) a estrutura geral do nucleossomo completo estabilizado por uma molécula de histona H1. Em (b), (c) e (d), os filamentos complementares de DNA são mostrados em marrom e verde e as histonas H2a, H2b, H3 e H4 são mostradas em amarelo, vermelho, azul e verde, respectivamente.

espécie para espécie e de um tipo de célula para outro. Ligadores tão pequenos quanto oito pares de nucleotídeos e tão grandes quanto 114 pares de nucleotídeos já foram relatados. Evidências sugerem que o nucleossomo completo (em oposição ao cerne do nucleossomo) exibe dois giros completos de super-hélice de DNA (um tamanho de 166 pares de nucleotídeos de DNA) na superfície do octâmero de histonas e a estabilização desta estrutura pela ligação de uma molécula de histona H1 (Fig. 9.24e).

O componente estrutural básico da cromatina eucariótica é o nucleossomo. Mas a estrutura de todos os nucleossomos é igual? Que papel(éis), se algum, a estrutura do nucleossomo tem na expressão gênica e na regulação da expressão gênica? Sabe-se que a estrutura dos nucleossomos nas regiões transcricionalmente ativas da cromatina difere da de nucleossomos em regiões transcricionalmente inativas. Quais são os detalhes desta correlação estrutura-função? As caudas de algumas moléculas de histona projetam-se do nucleossomo e estão acessíveis a enzimas que adicionam e removem grupos químicos tais como

metila ($-\text{CH}_3$) e grupos acetil ($-\text{C}(=\text{O})\text{CH}_3$). A adição

destes grupos muda o nível de expressão dos genes acondicionados em nucleossomos contendo histonas modificadas (veja Cap. 21).

Micrografias eletrônicas de cromossomos metafásicos isolados exibem massas de fibras muito helicoidizadas ou dobradas (Fig. 9.25). Essas **fibras de cromatina** têm um diâmetro médio de 30 nm. Quando as estruturas vistas por microscopia óptica e eletrônica durante os primeiros estágios da meiose são comparadas, torna-se claro que a microscopia óptica simplesmente permite ver tais regiões onde essas fibras de 30 nm são bem acondicionadas ou condensadas.

Qual a subestrutura da fibra de 30 nm vista em cromossomos mitóticos e meióticos? Embora biólogos não tenham uma resposta definitiva a esta pergunta, eles sabem que o DNA está enrolado como uma super-hélice ao redor de um octâmero de histonas para produzir o diâmetro de aproximadamente 10 nm do nucleossomo. *In vivo*, os nucleossomos estão provavelmente em justaposição direta um com o outro sem regiões ligadoras detectáveis. Assim, eles formarão uma fibra de nucleossomo de 10 nm (veja Fig. 9.23, em cima). Se essa fibra de 10 nm, por sua vez, está enrolada em uma super-hélice de ordem maior (um solenóide), pode ser gerada uma fibra de 30 nm. Embora os cientistas ainda não compreendam todos os detalhes da estrutura desta fibra de

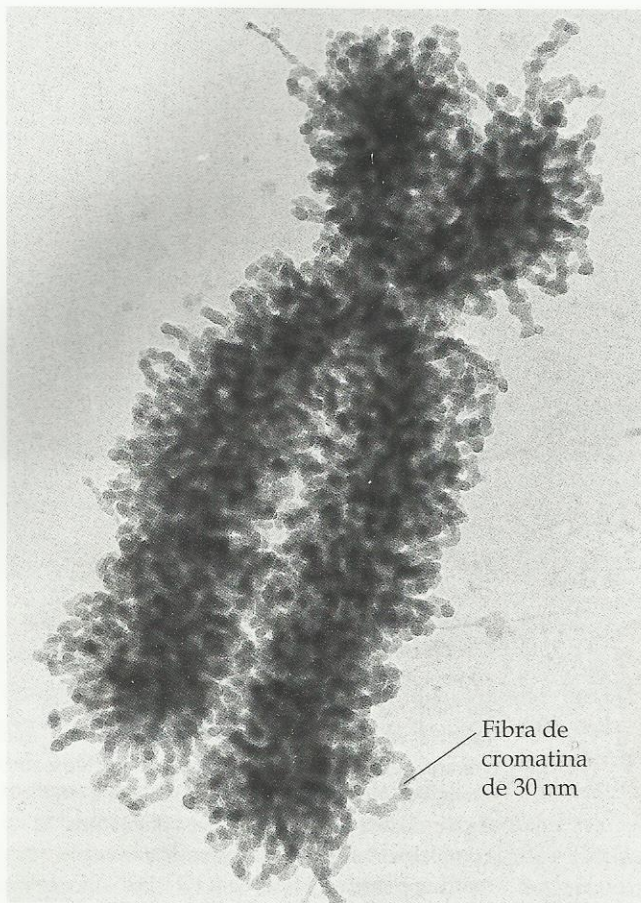


Fig. 9.25 ■ Micrografia eletrônica de um cromossomo metafásico humano mostrando a presença de fibras de cromatina de 30 nm. A evidência disponível indica que cada cromátide contém uma fibra de 30 nm altamente helicoidizada ou dobrada.

cromatina de 30 nm, existem boas evidências de que ela representa uma estrutura tipo solenóide tal como a mostrada na Fig. 9.26.

Os cromossomos metafásicos são os mais condensados dos cromossomos eucarióticos normais. Claramente, o papel desses cromossomos altamente condensados é organizar e compactar as moléculas gigantes de DNA de cromossomos eucarióticos em estruturas que irão facilitar sua segregação para os núcleos filhos sem que moléculas de DNA de cromossomos diferentes fiquem entrelaçadas e, como resultado, sejam rompidas durante a separação anafásica dos cromossomos filhos. Como vimos na seção anterior, a unidade estrutural básica do cromossomo metafásico é a fibra de cromatina de 30 nm. Entretanto, de que modo tais fibras de 30 nm ficam mais condensadas na estrutura metafásica observada? Infelizmente, ainda não existe uma resposta clara a esta pergunta. Existem evidências de que a estrutura geral de cromossomos metafásicos não é dependente de histonas. Micrografias eletrônicas de cromossomos metafásicos isolados dos quais foram removidas as histonas revelam um **arcabouço**, ou cerne, que é circundado por um enorme halo

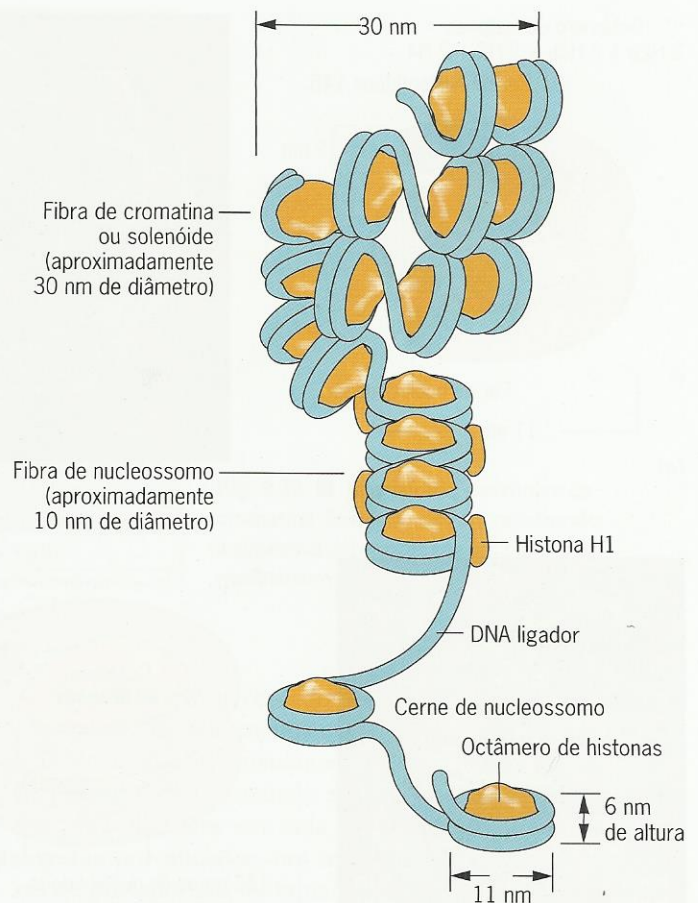


Fig. 9.26 ■ Diagrama do modelo solenóide da fibra de cromatina de 30 nm. A histona H1 parece estabilizar a fibra nucleossômica de 10 nm e contribuir para a formação da fibra de cromatina de 30 nm.

de DNA (Fig. 9.27). Esse arcabouço cromossômico deve ser composto de proteínas cromossômicas não histônicas. Note a ausência de pontas aparentes de moléculas de DNA na micrografia mostrada na Fig. 9.27. Este achado novamente apóia o conceito de uma molécula gigante de DNA por cromossomo.

Em resumo, são necessários pelo menos três níveis de condensação para acondicionar os 10^3 a 10^5 μm de DNA em um cromossomo eucariótico em uma estrutura metafásica com alguns microns de tamanho.

1. O primeiro nível de condensação envolve compactar o DNA como uma super-hélice negativa em nucleossomos, para produzir a fibra de cromatina interfásica de 10 nm de diâmetro. Isto claramente envolve um octâmero de moléculas de histona, duas de cada histona H2a, H2b, H3 e H4.
2. O segundo nível de condensação envolve um dobramento adicional ou super-hélice da fibra de nucleossomo de 10 nm, para produzir a fibra de cromatina de 30 nm característica dos cromossomos mitóticos e meióticos. A histona H1 está

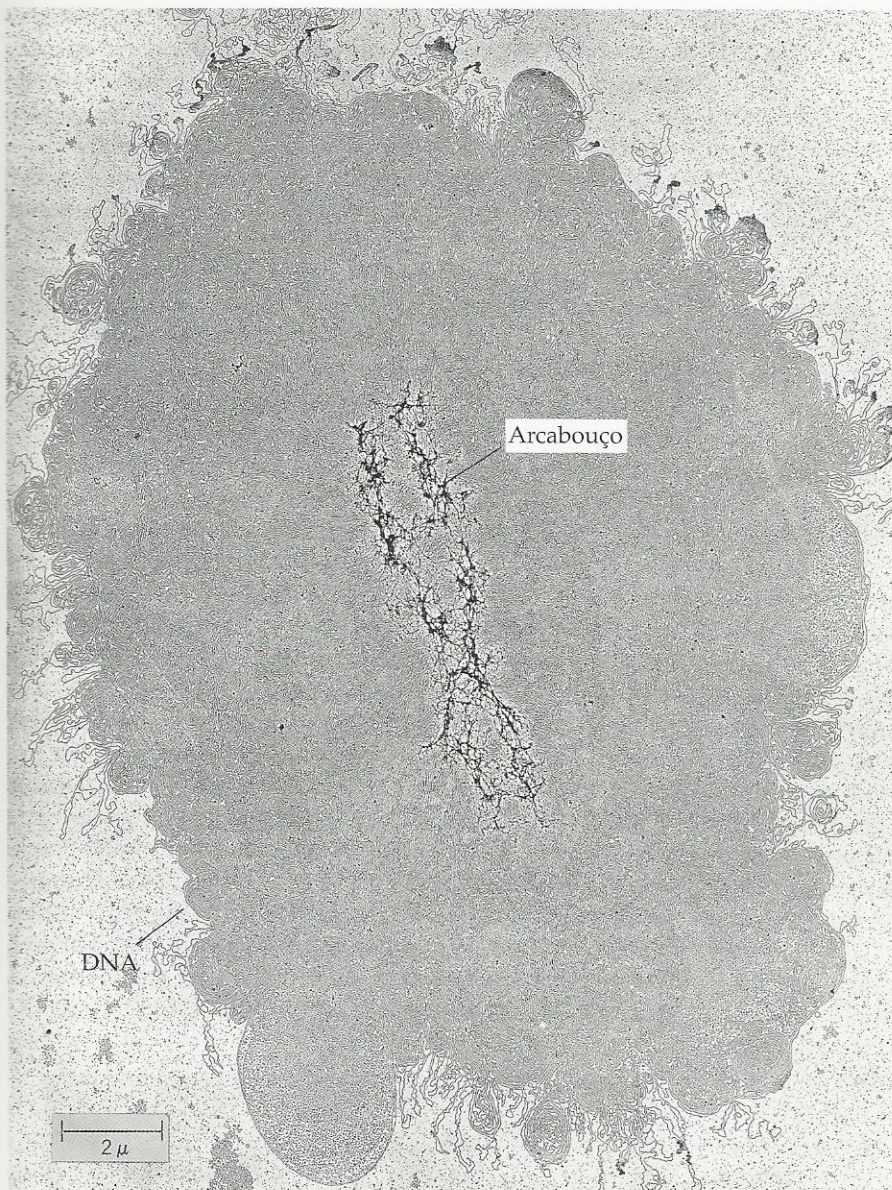


Fig. 9.27 ■ Micrografia eletrônica de um cromossomo humano metafásico do qual foram removidas as histonas. Um acúmulo enorme de DNA circunda um “arcabouço” central composto de proteínas cromossômicas não histônicas. Note que o arcabouço tem mais ou menos a mesma forma do cromossomo metafásico antes de remoção das histonas.

envolvida nesta super-helicoidização da fibra de 10 nm do nucleossomo para produzir a fibra de cromatina de 30 nm.

3. Finalmente, as proteínas cromossômicas não histônicas formam um arcabouço que está envolvido na condensação da fibra de cromatina de 30 nm nos cromossomos metafásicos muito compactados. Este terceiro nível de condensação parece envolver a separação de segmentos das moléculas gigantes de DNA presentes em cromossomos eucarióticos em domínios super-helicoidizados independentemente, ou alças (Fig. 9.27). O mecanismo pelo qual ocorre este terceiro nível de condensação não é conhecido.

Centrômeros e Telômeros

Como discutimos no Cap. 2, os dois cromossomos homólogos (cada um contendo duas cromátides irmãs) de cada par

cromossômico separam-se para pólos opostos do fuso meiótico durante a anáfase I da meiose. Similarmente, durante a anáfase II da meiose e a anáfase da mitose, as cromátides irmãs de cada cromossomo movem-se para pólos opostos do fuso e se tornam cromossomos filhos. Esses movimentos anafásicos dependem da ligação de microtúbulos a regiões específicas dos cromossomos, os centrômeros. Como todos os centrômeros desempenham a mesma função básica, não surpreende que os centrômeros de cromossomos diferentes de uma espécie contenham componentes estruturais similares.

O centrômero de um cromossomo metafásico (Fig. 9.28) em geral pode ser reconhecido como uma região de constrição (veja Fig. 9.25) onde o cromossomo não parece ter-se duplicado. De fato, a produção de dois centrômeros funcionais a partir de um centrômero parental é uma etapa fundamental na transição da metáfase para a anáfase, e um

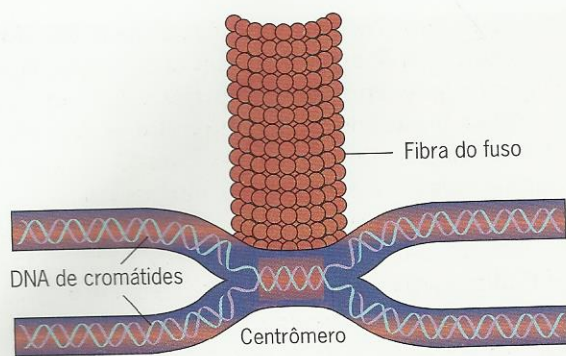


Fig. 9.28 ■ Modelo da estrutura do centrômero em um cromossomo metafásico. As fibras do fuso, que se ligam a centrômeros, são responsáveis pela separação de cromossomos homólogos durante a anáfase I da meiose e os cromossomos da prole (derivados de cromátides) durante a anáfase II da meiose e a anáfase da mitose (Cap. 1).

centrômero funcional deve estar presente em cada cromossomo filho para evitar os efeitos deletérios de não-disjunção. Fragmentos cromossômicos acêntricos geralmente são perdidos durante divisões mitóticas e meióticas.

Os centrômeros (regiões *CEN*) de todos os cromossomos da levedura do pão, *S. cerevisiae*, foram isolados e caracterizados. As regiões *CEN* de cromossomos diferentes são intercambiáveis: a substituição da região *CEN* de um cromossomo pela região *CEN* de outro cromossomo não tem efeito detectável na célula hospedeira ou na sua capacidade de sofrer uma divisão celular normal. Estudos moleculares mostraram que um centrômero funcional de *S. cerevisiae* tem de 110 a 120 pares de nucleotídeos de tamanho e possui três regiões essenciais (Fig. 9.29). As regiões I e III são pequenas seqüências limites conservadas, e a região II é um segmento central rico em A-T (>90% A-T) com cerca de 90 pares de nucleotídeos de tamanho. O tamanho e a natureza rica em A-T da região II provavelmente são mais importantes do que sua seqüência de nucleotídeos, enquanto as regiões I e III contêm seqüências específicas que servem como sítios de ligação para proteínas envolvidas na ligação de fibras do fuso.

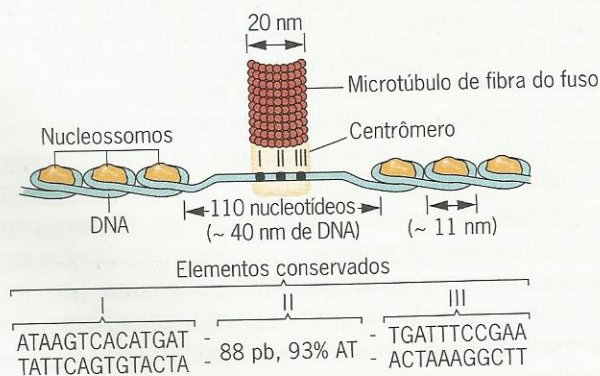


Fig. 9.29 ■ Diagrama da estrutura conservada dos centrômeros em *Saccharomyces cerevisiae* (em cima) e a seqüência da região *CEN* do cromossomo 3 desta espécie (embaixo).

Os centrômeros de eucariontes multicelulares são muito maiores e mais complexos do que os da levedura do pão, mas sua estrutura real ainda é incerta. Os centrômeros de plantas e animais superiores contêm grandes quantidades de seqüências de DNA que são repetidas muitas vezes, frequentemente em longas disposições em tandem. Outras seqüências de DNA são geralmente encontradas inseridas nestas disposições em tandem. Cada centrômero de cromossomos humanos, por exemplo, contém de 5.000 a 15.000 cópias de uma seqüência de 171 pares de bases de tamanho chamada seqüência satélite alfa (às vezes "alfóide") (Fig. 9.30). (As seqüências satélites formam bandas "satélites" durante a centrifugação em gradiente de densidade. Veja Enfoque Técnico: Técnicas de Centrifugação no Cap. 10.) Huntington Willard e colaboradores mostraram que um segmento de 450.000 pares de bases do centrômero do cromossomo X humano é suficiente para funcionamento do centrômero. Esse segmento consiste principalmente em seqüências alfa satélites, mas contém sítios de ligação de proteína centromérica (CENP) intercalados chamados boxes CENP-B. Ambos os componentes são essenciais para o funcionamento do centrômero.

Há várias décadas, sabemos que os **telômeros** (dos termos gregos *telos* e *meros*, que significam "fim" e "parte"), ou extremidades de cromossomos eucarióticos, têm propriedades únicas. Hermann J. Muller, que introduziu o termo *telômero* em 1938, demonstrou que os cromossomos de drosófilas sem pontas naturais, produzidas por quebras cromossômicas com raios X, não eram transmitidos para a prole. Em um estudo clássico de cromossomos de milho, Barbara McClintock (veja Enfoque Técnico: Barbara McClintock, a Descobridora dos Elementos de Transposição, no Cap. 18)



Fig. 9.30 ■ A localização das seqüências de DNA alfa satélite (amarelas) nos centrômeros de cromossomos humanos (vermelhos). Veja Enfoque Técnico: Hibridização *In Situ*.

demonstrou que novas pontas de cromossomos quebrados são adesivas e tendem a se fundir umas com as outras. Em contraste, as pontas naturais de cromossomos normais (não quebradas) são estáveis e não mostram tendência em se fundir com outras pontas quebradas ou nativas. Os resultados de McClintock indicaram que telômeros devem ter estruturas especiais diferentes das pontas produzidas por quebra de cromossomos.

Outro motivo para postular que os telômeros têm estruturas únicas é que os mecanismos conhecidos de replicação de moléculas lineares de DNA não permitem duplicação de ambos os filamentos de DNA nas pontas das moléculas (Cap. 10). Assim, os telômeros devem ter estruturas únicas que facilitam sua replicação, ou deve haver alguma enzima especial de replicação que resolve este enigma. Qualquer que seja sua estrutura, os telômeros devem desempenhar pelo menos três funções importantes. Eles devem (1) impedir que desoxirribonucleases degradem as pontas das moléculas lineares de DNA, (2) evitar fusão das pontas com outras moléculas de DNA e (3) facilitar a replicação das pontas das moléculas lineares de DNA sem perda de material.

Os telômeros de cromossomos eucarióticos têm estruturas únicas que incluem seqüências curtas de nucleotídeos presentes em repetições em tandem. Embora as seqüências variem um pouco em diferentes espécies, a unidade básica repetida tem o padrão 5'T₁₋₄A₀₋₁G₁₋₈-3' em quase todas as espécies. Por exemplo, a seqüência repetida em humanos e outros vertebrados é TTAGGG, a do protozoário *Tetrahymena thermophila* é TTGGGG e a da planta *Arabidopsis thaliana* é TTTAGGG. Na maioria das espécies, as seqüências repetitivas adicionais de DNA estão presentes adjacentes a telômeros. Elas são chamadas seqüências associadas a telômeros.

Em vertebrados, a repetição TTAGGG é altamente conservada; ela foi identificada em mais de 100 espécies, incluindo mamíferos, aves, répteis, anfíbios e peixes. Várias cópias desta unidade básica repetida em telômeros variam de espécie para espécie, de cromossomo para cromossomo dentro de uma espécie e até mesmo no mesmo cromossomo em diferentes tipos celulares. Em células somáticas humanas normais (não cancerosas), os telômeros geralmente contêm de 500 a 3.000 repetições TTAGGG e gradualmente se encurtam com a idade. Em contraste, os telômeros das células da linhagem germinativa e das células cancerosas não se encurtam com a idade (veja Tamanho do Telômero e Envelhecimento em Humanos, no Cap. 10).

Os telômeros de algumas espécies não são compostos de curtas repetições em tandem do tipo descrito. Em *D. melanogaster*, por exemplo, os telômeros são compostos de duas seqüências especializadas de DNA que podem mover-se de um local do genoma para outros locais. Devido a sua mobilidade, tais seqüências são chamadas *elementos genéticos transponíveis* (veja Cap. 18).

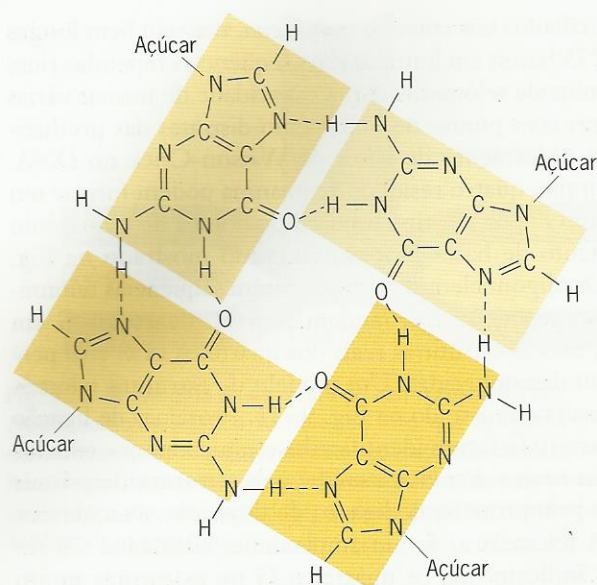
A maioria dos telômeros terminam com uma região unifilar rica em G do filamento de DNA com a ponta 3' (a chamada protruída 3'). Essas projeções são curtas (12 a 16 ba-

ses) em ciliados tais como *Tetrahymena*, mas são bem longas (125 a 275 bases) em humanos. As seqüências repetidas ricas em guanina de telômeros têm a capacidade de formar várias estruturas com pontes de hidrogênio distintas das produzidas pelo pareamento de bases de Watson-Crick no DNA. Por exemplo, quatro resíduos de guanina podem formar um "quarteto G" por um tipo especial de pontes de hidrogênio (pareamento de bases Hoogsteen) como mostrado na Fig. 9.31a. Os oligonucleotídeos que contêm seqüências teloméricas com repetições em tandem formam quartetos G em solução, mas as estruturas reais dos quartetos G *in vivo* permanecem desconhecidas. Um modelo de estrutura telomérica humana é mostrado na Fig. 9.31b. Proteínas de ligação com quarteto G foram identificadas e estudadas nos ciliados *Oxytricha nova* e *Euplotes crassus*, e a estrutura de cristais formada pela proteína de ligação de *Oxytricha* e seqüências de DNA teloméricas foram amplamente estudadas. Os resultados indicam que os quartetos G ou estruturas muito similares têm um importante papel em estabilizar as pontas de cromossomos lineares. Em humanos, foi demonstrado que uma proteína chamada **TRF-2** (*Telomere Repeat-binding Factor 2*) evita a fusão ponta a ponta de cromossomos.

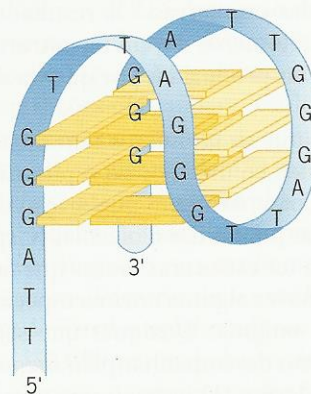
Recentemente, foi mostrado que os telômeros de humanos e algumas outras espécies formam estruturas chamadas **alças-t**, nas quais o único filamento na ponta 3' invade uma repetição telomérica antecedente e faz par com o filamento complementar, deslocando o filamento equivalente (Fig. 9.31c). O filamento deslocado é protegido de degradação por ser revestido com uma proteína chamada POT-1 (**P**rotection **O**f **T**elomeres-1 – proteção de telômeros-1). Note que a estrutura em alça-t protege bem a ponta livre da molécula de DNA. Dois complexos protéicos específicos de telômeros, TRF-1 e TRF-2, estão associados a essas alças-t. Ambos os complexos contêm várias proteínas conhecidas que não são identificadas aqui por simplificação. Uma destas proteínas é uma enzima que desenrola o DNA e pode perturbar os pares de bases Hoogsteen em quartetos G, permitindo que o terminal 3' invada uma região telomérica antecedente e forme uma alça-t. Hoje em dia, alças-t foram identificadas nos telômeros de vertebrados, no ciliado *Oxytricha fallax*, no protozoário *Trypanosoma brucei* e na planta *Pisum sativum* (ervilhas). Assim, eles podem ser componentes importantes dos telômeros da maioria das espécies.

■ PONTOS IMPORTANTES

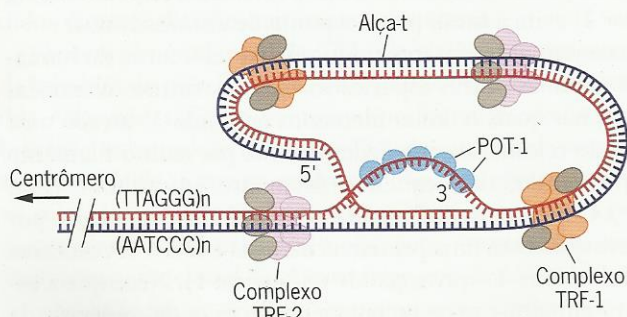
- O cromossomo eucariótico contém uma molécula gigante de DNA compactada em contos elípticos de 10 nm chamadas nucleossomos.
- Os cromossomos condensados que estão presentes na mitose e na meiose são compostos de fibras de cromatina de 30 nm.
- Na metáfase, as fibras de 30 nm são segregadas em domínios por arcabouços compostos de proteínas cromossômicas não histônicas.
- Os centrômeros (regiões de ligação das fibras do fuso) e os telômeros (pontas) de cromossomos têm estruturas únicas que facilitam suas funções.



(a) Estrutura de um quarteto G



(b) Estrutura unimolecular de quarteto G



(c) Modelo de um telômero humano estabilizado pela formação de uma alça em t.

Fig. 9.31 ■ (a) Estrutura do quarteto G formado em solução por oligonucleotídeos ricos em G tipo telômero, (b) um modelo em grampo unimolecular da estrutura do telômero humano e (c) um telômero humano estabilizado pela formação de uma alça-t, na qual o 3'-terminal invade uma repetição antecedente de telômero e pareia com o filamento complementar. TRF-1 e TRF-2 são os fatores 1 e 2 de ligação de repetições do telômero; ambos são complexos contendo várias proteínas. A proteína POT-1 (proteína 1 de proteção dos telômeros) reveste o filamento único de DNA deslocado pela invasão do terminal 3' do DNA telomérico.

GENOMAS EUCARIÓTICOS: SEQÜÊNCIAS REPETIDAS DE DNA

Muitas das seqüências de DNA nos genomas eucarióticos estão presentes em várias cópias, às vezes até um milhão.

Os cromossomos dos procariontes contêm quase que exclusivamente moléculas de DNA com seqüências únicas de pares de bases (não repetidas); isto é, cada gene (com algumas exceções) está presente apenas uma vez no genoma. Se as moléculas de DNA em cromossomos procarióticos são quebradas em vários segmentos pequenos, cada fragmento conterá uma seqüência diferente de pares de bases. Os cromossomos de eucariontes são muito mais complexos quanto a isto. Algumas seqüências de bases são repetidas muitas vezes no complemento cromossômico haplóide, ocasionalmente até um milhão de vezes. O DNA contendo tais seqüências repetidas, chamado **DNA repetitivo**, em geral representa um componente principal (20 a 50%) do genoma eucariótico.

A primeira evidência de DNA repetitivo veio de estudos sobre centrifugação de DNA eucariótico. Quando o DNA de um procarionte, tal como *E. coli* é isolado, fragmentado e centrifugado em altas velocidades por longos períodos de tempo em uma solução de cloreto de cério 6 M (CsCl), o DNA formará uma única banda no tubo de centrífuga na posição em que sua densidade é igual à densidade da solução de CsCl (Cap. 10). Para *E. coli*, esta banda será formada na posição em que a densidade do CsCl é igual à densidade do DNA contendo cerca de 50% de pares A-T e 50% G-C. A densidade do DNA aumenta com o aumento do conteúdo de G-C. As pontes de hidrogênio extras em um par de bases G-C resulta em uma associação mais forte entre as bases e assim em uma densidade maior do que a dos pares de bases A-T. A centrifugação do DNA de eucariontes para condições de equilíbrio em tais soluções de CsCl geralmente revela a presença de uma banda principal grande de DNA e uma a várias bandas pequenas. Essas pequenas bandas de DNA são chamadas **bandas satélites** (da palavra latina *satelles*, que significa "anexa" ou "subordinada") e os DNA nestas bandas geralmente são chamados **DNA satélites**. Por exemplo,

o genoma de *Drosophila virilis*, um parente distante de *Drosophila melanogaster*, contém três DNA satélites distintos, cada um composto de uma sequência repetida de sete pares de bases. Outros DNA satélites em eucariontes têm longas sequências repetitivas.

Detecção de Sequências Repetidas: Cinética de Renaturação de DNA

O primeiro detalhamento da frequência e da complexidade das sequências repetidas de DNA em eucariontes resultou de estudos das velocidades de renaturação do DNA. Os dois filamentos de uma dupla hélice de DNA são mantidos juntos por um grande número de pontes de hidrogênio relativamente fracas entre bases complementares. Quando moléculas de DNA em solução aquosa são aquecidas a quase 100°C, essas ligações são quebradas e os filamentos complementares de DNA se separam. Este processo é chamado **desnaturação**.

Como os pares de bases G-C contêm três pontes de hidrogênio e os pares A-T contêm duas pontes de hidrogênio, é necessário mais energia para romper o pareamento de bases em moléculas de DNA ricas em G-C do que as duplas hélices ricas em A-T. Se os filamentos únicos complementares de DNA são resfriados lentamente em condições corretas, as sequências complementares de bases se encontrarão e reconstituirão as duplas hélices com pares de bases. Essa reconstituição de duplas hélices a partir de filamentos únicos complementares de DNA é chamada **renaturação**. Análises da cinética de renaturação de DNA de organismos eucarióticos produziram muitas informações sobre os tipos de sequências repetidas de DNA presentes nos genomas de eucariontes.

Consideremos uma longa molécula de DNA sem sequências repetidas (por exemplo, uma molécula de DNA de um cromossomo procariótico). Se tal molécula for fracionada em fragmentos de um determinado tamanho, digamos 400 pares de nucleotídeos, for desnaturada e for deixada renaturar-se sob condições apropriadas, sua velocidade de renaturação dependerá (1) da concentração de DNA na solução e (2) da complexidade do DNA, isto é, do número de fragmentos diferentes de 400 pares de bases. Estes dois fatores determinarão a probabilidade de que dois filamentos únicos complementares de DNA colidam e ocorra renaturação. O efeito da concentração de DNA é bem óbvio, pois, quanto mais alta a concentração de filamentos únicos na solução, maior a chance de uma colisão entre quaisquer dois deles.

A complexidade do DNA é o tamanho total de sequências de pares de nucleotídeos não repetidos no genoma. Em uma dupla hélice com poli A em um filamento e poli T no outro filamento, os filamentos complementares de cada fragmento serão iguais. Assim, todas as colisões entre filamentos únicos de DNA derivados de filamentos complementares resultarão em renaturação. À medida que a complexidade do DNA aumenta, a proporção de colisões aleatórias entre filamentos únicos de DNA que estão entre filamentos únicos com-

plementares diminuirá, e a taxa de renaturação diminuirá concordantemente. Considere genomas que contêm 4, 400, 4.000 e 40.000 genes diferentes. Em uma solução contendo 100 µg de DNA por ml, a concentração de qualquer gene será de 25, 0,25, 0,025 e 0,0025 µg/ml, respectivamente, para os quatro genomas. Como resultado, colisões aleatórias entre os filamentos complementares derivados de qualquer gene diminuirão de frequência à medida que a complexidade do genoma (o número de genes diferentes) aumenta. Assim, espera-se que sequências de grandes genomas renaturem-se lentamente.

Surpreendentemente, algumas sequências de DNA nos genomas complexos de eucariontes renaturam-se muito rapidamente. Essas sequências são repetidas muitas vezes no genoma, ocasionalmente um milhão de vezes ou mais. Uma sequência que é repetida no genoma estará presente em uma concentração mais alta que uma sequência de cópia única durante renaturação, aumentando assim sua velocidade de renaturação. De fato, a velocidade de renaturação do DNA é diretamente proporcional ao número de cópias – quanto maior o número de cópias, mais rápida a velocidade e menor o tempo necessário para renaturação. Análises matemáticas das velocidades de renaturação de genomas eucarióticos forneceram evidências iniciais da presença de classes diferentes de sequências repetidas de DNA, ou DNA repetitivo. Os recentes projetos de sequenciamento do genoma deram informações adicionais sobre os tipos diferentes de sequências de DNA repetitivo em genomas eucarióticos. As localizações de sequências diferentes de DNA em cromossomos podem ser determinadas diretamente por procedimentos similares aos experimentos de renaturação já descritos. Com este procedimento, chamado “hibridização *in situ*”, filamentos marcados de DNA formam duplas hélices com DNA desnaturado ainda presente em cromossomos (veja Enfoque Técnico: Hibridização *In Situ*).

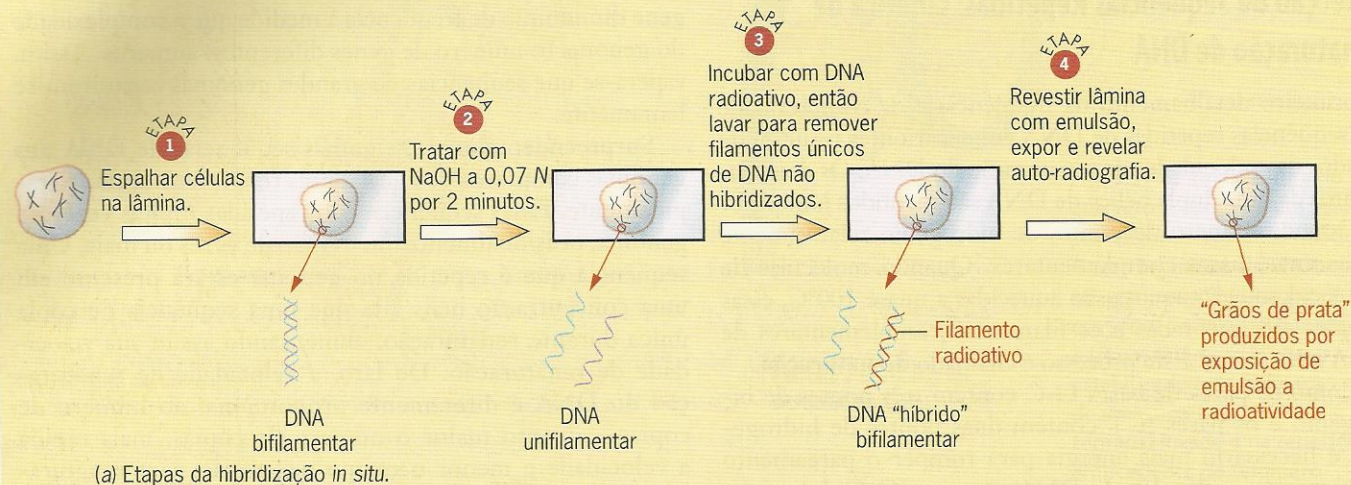
Sequências de DNA de Cópia Única e Repetitivas em Genomas Eucarióticos

Estudos de cinética de renaturação do DNA foram feitos em muitos genomas eucarióticos diferentes, revelando a grande variabilidade nas quantidades e nos tipos de DNA repetitivo. Alguns eucariontes contêm relativamente pouco DNA repetitivo; no verme *Caenorhabditis elegans*, apenas 17% do genoma consiste em sequências repetidas. Outros eucariontes contêm mais DNA repetitivo que o DNA de cópia única; 72% do DNA da planta do tabaco, *Nicotiana tabacum*, é composto de sequências repetidas. Pelo menos 50% do genoma humano consiste em sequências repetidas de DNA. As porções do genoma humano que ainda não foram sequenciadas provavelmente consistem em sequências de DNA repetidas. Assim, a quantidade de DNA repetitivo provavelmente será maior que 50% quando a sequência estiver completa.

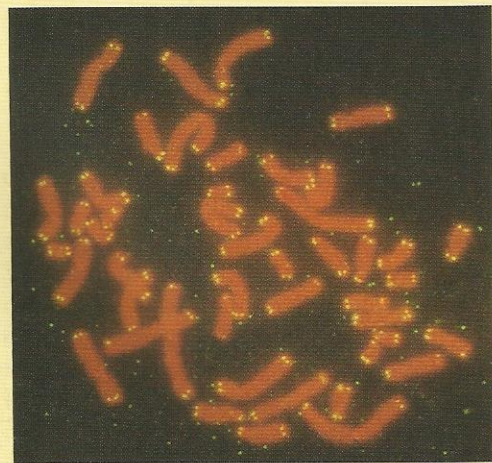
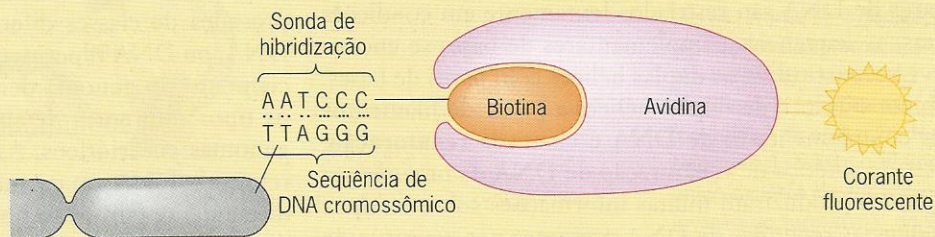
ENFOQUE TÉCNICO: Hibridização *In Situ*

Em 1969, Mary Lou Pardue e Joseph Gall desenvolveram um procedimento pelo qual eles podiam helicoidizar filamentos isolados

de DNA radioativo com filamentos complementares de DNA em cromossomos sobre lâminas de vidro. Usando este procedimento,



(b) Auto-radiografia mostrando localizações cromossômicas de seqüências de DNA satélite de camundongo.



(d) Telômeros humanos visualizados usando sondas fluorescentes e hibridização *in situ*.

Fig. 1 ■ Localização de seqüências repetidas de DNA em cromossomos por hibridização *in situ* feita com sondas radioativas (a e b) ou sondas fluorescentes (c e d). O procedimento de hibridização *in situ* desenvolvido por Pardue e Gall é mostrado em (a), e uma de suas auto-radiografias demonstrando a presença da seqüência do DNA satélite de camundongo na heterocromatina centromérica é mostrada em (b). O uso de corantes fluorescentes para localizar a seqüência repetida TTAGGG nos telômeros de cromossomos humanos é ilustrado em (c), e uma fotomicrografia demonstrando sua localização centromérica é mostrada em (d).



ENFOQUE TÉCNICO: Hibridização *In Situ* (continuação)

chamado **hibridização *in situ***, Pardue e Gall foram capazes de determinar as localizações cromossômicas de seqüências de DNA repetitivo. (O termo latino *in situ* significa "em seu lugar original"; hibridização é a formação de moléculas dúplices "híbridas" pela helicoidização de filamentos complementares ou parcialmente complementares de DNA ou RNA.) A hibridização *in situ* clássica consiste em espalhar cromossomos mitóticos em uma lâmina de vidro (veja Fig. 6.1), desnaturar o DNA nos cromossomos por exposição a álcali (NaOH a 0,07 N) por alguns minutos, lavar com tampão para remover a solução alcalina, incubar a lâmina em solução de hibridização contendo cópias radioativas da seqüência de nucleotídeos de interesse, remover os filamentos radioativos que não se hibridizaram com seqüências complementares nos cromossomos, expor a lâmina a uma emulsão fotográfica que é sensível a radioatividade de baixa energia, revelar a auto-radiografia e superpor a auto-radiografia e uma foto dos cromossomos (Fig. 1a e b).

Um dos primeiros experimentos de hibridização *in situ* que Pardue e Gall fizeram demonstrou que a seqüência de DNA satélite do camundongo está localizada em regiões heterocromáticas flanqueadoras dos centrômeros dos cromossomos de camundongo (Fig. 1b). O genoma do camundongo contém cerca de 10^6 cópias desta seqüência de DNA satélite, que tem cerca de 400 pares de nucleotídeos de tamanho e constitui cerca de 10% do genoma de camundongo. Estudos similares foram subsequentemente feitos com DNA satélites de várias outras espécies, e

tais seqüências de DNA repetitivo estão geralmente situadas na heterocromatina centromérica ou adjacente aos telômeros. Uma seqüência de DNA repetitivo pode ser identificada como DNA satélite apenas se a seqüência tiver uma composição de bases suficientemente diferente da do DNA da banda principal para produzir uma banda distinta durante centrifugação em gradiente de densidade. Portanto, a centrifugação não pode ser usada para identificar todas as seqüências de DNA repetitivo. Seqüências de DNA satélite geralmente não são expressas; isto é, elas não codificam RNA ou proteínas.

Hoje em dia, experimentos de hibridização *in situ* são em geral feitos usando-se sondas de hibridização que são ligadas a corantes fluorescentes ou anticorpos marcados com compostos fluorescentes (Fig. 1c e d). Em um protocolo, sondas de hibridização de DNA ou RNA são ligadas à vitamina biotina, que é ligada com alta afinidade pela proteína do ovo avidina (Fig. 1c). Usando-se avidina covalentemente ligada a um corante fluorescente, a localização cromossômica da sonda hibridizada pode ser detectada pela fluorescência do corante. Este procedimento, chamado FISH (hibridização *in situ* com fluorescência), foi usado para demonstrar a presença da seqüência repetida TTAGGG nos telômeros de cromossomos humanos (Fig. 1d). O procedimento FISH é muito sensível e pode ser usado para detectar as localizações de seqüências de cópia única em cromossomos interfásicos e mitóticos humanos (veja Fig. 16.6b).

As seqüências de DNA presentes em eucariontes são comumente agrupadas em três classes:

1. **Seqüências únicas, ou de uma só cópia, de DNA** – 1 a 10 cópias por genoma.
2. **Seqüências de DNA moderadamente repetitivas** – 10 a 10^5 cópias por genoma.
3. **Seqüências de DNA altamente repetitivas** – mais de 10^5 cópias por genoma.

Estas classes, especialmente a de DNA moderadamente repetitivo, são bem amplas e heterogêneas em termos de número de cópias e tipos de seqüências.

Seqüências Únicas ou de Cópia Única

As seqüências de DNA de cópia única constituem cerca de 30 a 85% dos genomas de eucariontes, enquanto representam quase 100% das seqüências em procariontes. Cerca de 90, 83, 70, 50, e 28% das seqüências de DNA nos genomas da planta *Arabidopsis*, de *Caenorhabditis*, de *Drosophila*, de humanos e de plantas do tabaco, respectivamente, estão presentes em uma ou poucas cópias. A maioria dos genes que codificam proteínas estão presentes neste componente do genoma.

Além disso, muitas das seqüências que regulam a expressão dos genes estão presentes nesta fração do genoma. Assim, as seqüências de DNA de cópia única estocam uma grande proporção da informação genética de um organismo.

Seqüências de DNA Moderadamente Repetitivas

Seqüências moderadamente repetitivas constituem desde 5% até mais de 80% dos genomas de eucariontes. Em *Arabidopsis*, *Caenorhabditis*, *Drosophila*, humanos e plantas do tabaco, as seqüências moderadamente repetitivas representam cerca de 10, 14, 12, 45 e 65%, respectivamente, das seqüências no genoma. As seqüências de DNA moderadamente repetitivas pertencem à mesma classe de seqüência apenas em termos de número de cópias. Em outros aspectos, elas são muito heterogêneas, incluindo uma disposição diferente de seqüências repetidas. A maioria das seqüências moderadamente repetitivas estão presentes em 10 a 1.000 cópias por genoma, mas algumas ocorrem em números mais altos de cópias. Muitas das seqüências moderadamente repetitivas são distribuídas pelo genoma intercaladas com seqüências de cópia única. Este padrão intercalado é ilustrado para uma curta região heterocromática do cromossomo 5 de *Arabidopsis* na Fig. 9.32.

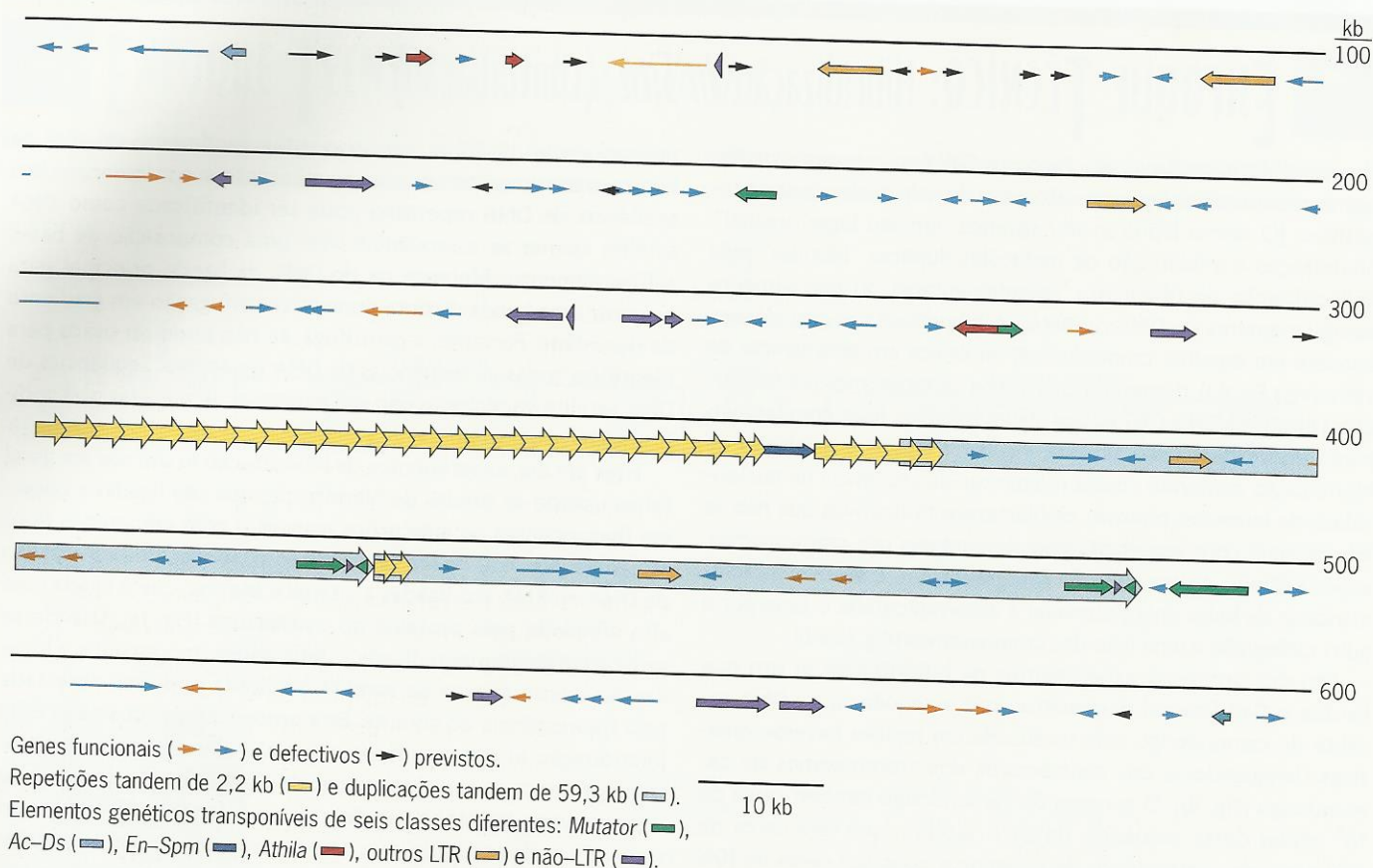


Fig. 9.32 ■ Sequências de cópia única e repetitivas intercaladas em uma protuberância heterocromática (chamada hk5L) perto do centrômero do cromossomo 5 na pequena crucifera *A. thaliana*. O DNA repetitivo intercalado inclui tanto repetições em tandem quanto 26 elementos transponíveis.

Algumas proteínas, tais como as proteínas ribossômicas e as proteínas musculares actina e miosina, são necessárias em grandes quantidades, e cada uma delas é codificada por vários genes. Esses genes de várias cópias ou famílias de genes em geral estão presentes na fração moderadamente repetitiva do genoma. Os genes que especificam RNA ribossômicos também são moderadamente repetitivos pois as células precisam de grandes quantidades de RNA ribossômico para produzir os ribossomos necessários para a síntese de proteínas.

A mais prevalente das sequências moderadamente repetitivas são os **elementos genéticos transponíveis**, sequências de DNA que podem mover-se de um local em um cromossomo para outro ou mesmo para um cromossomo diferente (Cap. 18), ou sequências inativas derivadas de elementos de transposição. Em *D. melanogaster*, cerca de 90 famílias diferentes de elementos de transposição foram caracterizadas e receberam nomes curiosos tais como *hobo*, *pogo* e *gypsy* que sugerem sua mobilidade. Uma proporção muito maior do genoma humano é derivada de elementos transponíveis. De acordo com os dados do International Human Genome Sequencing Consortium, cerca de 44% do genoma humano consiste em elementos transponíveis ou sequências deriva-

das deles. Até 80% do genoma do milho pode consistir em elementos genéticos transponíveis. Embora a maioria desses elementos genéticos transponíveis sejam moderadamente repetitivos, alguns são altamente repetitivos.

Sequências de DNA Altamente Repetitivas

Sequências altamente repetitivas constituem apenas 3% do genoma de *Caenorhabditis*, 18% do genoma de *D. melanogaster* e 45% do genoma de *Drosophila virilis*. Em muitas espécies, grande parte do DNA repetitivo está presente em telômeros e centrômeros (veja Enfoque Técnico: Hibridização *In Situ* e Fig. 9.33). Em *D. virilis*, três sequências, cada uma com sete pares de bases de tamanho – 5'-ATAAACT-3', 5'-ACAAACT-3' e 5'-ACAAATT-3' – formam bandas satélites distintas quando analisadas por centrifugação em gradiente de densidade (Cap. 10). Entretanto, a ausência de sequências de DNA altamente repetitivas similares em espécies proximamente relacionadas de *Drosophila* sugere que tais sequências não têm papéis essenciais no crescimento e no desenvolvimento normais. Em várias espécies de caranguejos terrestres, até 25% do DNA consiste em uma sequência altamente repetitiva que contém 97% de pares de

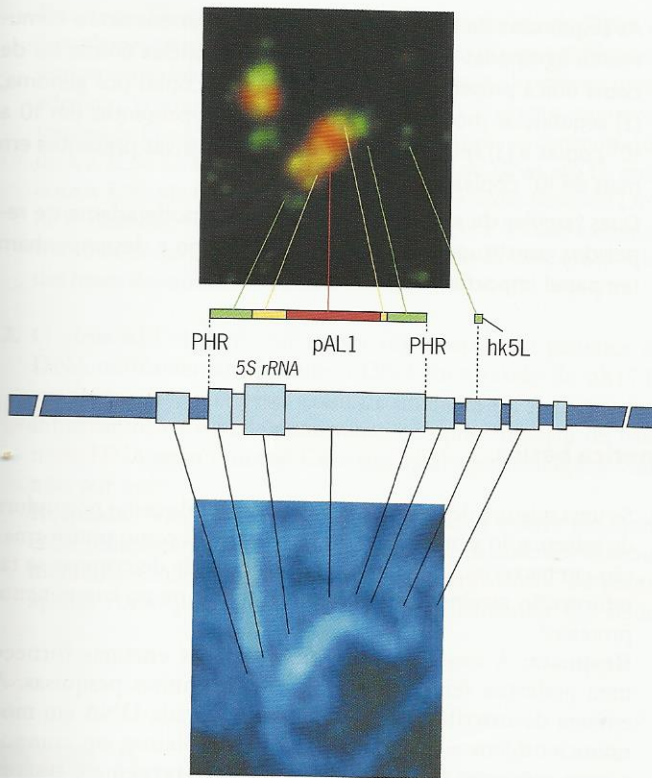


Fig. 9.33 ■ (Em cima) Localizações de seqüências de DNA repetitivas nas regiões heterocromática pericentromérica e centromérica (PHR) do cromossomo 5 de *Arabidopsis*. Hibridização *in situ* com fluorescência (veja Enfoque Técnico: Hibridização *In Situ*) foi usada para determinar as localizações de (1) repetições de 180 pb similares às seqüências chamadas repetições alfa satélites nos centrômeros de primatas (fluorescência vermelha na região pAL1), (2) elementos transponíveis (verde) *Athila* e (3) repetições em tandem dos genes de RNA 5S ribossômico (amarelo). Uma protuberância heterocromática chamada hk5L (veja Fig. 9.32) é visível como uma ilha verde-clara à direita. (Embaixo) A região centromérica do cromossomo 5 no estágio de paquíteno da prófase corada para mostrar segmentos mais condensados de heterocromatina.

bases A-T com uma seqüência A-T alternada. Claramente, estas seqüências não codificam proteínas. Na verdade, nenhuma seqüência de DNA altamente repetitiva estudada até hoje foi transcrita.

Quais são então as funções destas seqüências de DNA altamente repetitivas? No momento, não há uma resposta clara a esta pergunta. Funções postuladas para o DNA altamente repetitivo incluem: (1) papéis de organização ou estruturais em cromossomos; (2) envolvimento em pareamento cromossômico durante meiose; (3) envolvimento em *crossing-over* ou recombinação (Cap. 13); (4) proteção de importantes genes estruturais, como genes de histonas, de rRNA ou de proteínas ribossômicas; (5) um repositório de seqüências de DNA não essenciais para uso na futura evolução da espécie; e (6) sem nenhuma função, apenas DNA sucatas que é levado juntamente com os processos de replicação

e segregação de cromossomos. A validade de qualquer destes papéis postulados permanece por ser estabelecida.

Seqüências Repetidas no Genoma Humano

Em humanos, como em outros eucariontes, grande parte do DNA moderadamente repetitivo disperso consiste em famílias de elementos de transposição. A seqüência do genoma humano mostra que pelo menos 44% do genoma é derivado de elementos de transposição. Além disso, grande parte do componente de cópia única do DNA do genoma pode ser derivada de cópias antigas de elementos de transposição que divergiram muito para serem reconhecidas como derivadas de elementos transponíveis. A maioria dos elementos transponíveis no genoma humano e nos genomas de outros mamíferos pertencem a quatro classes: **LINE** (longos elementos nucleares intercalados), **SINE** (curtos elementos nucleares intercalados), elementos transponíveis contendo **longas repetições terminais (LTR)** e **transposons de DNA** (veja Cap. 18).

Os elementos transponíveis LINE funcionais têm tamanho médio de cerca de 6 kb; eles estão dispersos pelo genoma humano. O único membro funcional da família LINE humana é o elemento transponível *L1* (Cap. 18), que está presente em cerca de 516.000 cópias por genoma. Os elementos *L2* e *L3* são derivados inativos de *L1*. Juntos, *L1*, *L2* e *L3* constituem cerca de 20% do genoma humano.

Os elementos SINE têm apenas de 100 a 400 pares de bases de tamanho. O único elemento SINE ativo no genoma humano é chamado **Alu**. (O nome é derivado da enzima que corta o DNA em uma seqüência específica situada dentro do elemento. Veja Designação das Enzimas de Restrição, no Cap. 15.) Existem aproximadamente 1.090.000 seqüências Alu no genoma humano haplóide. Elas constituem cerca de 10% do genoma. Dois elementos SINE inativos constituem outros 3% do genoma humano.

Elementos transponíveis que contêm longas repetições terminais constituem cerca de 8% do genoma humano. Eles têm estrutura bem diversa e, em muitos casos, consistem em apenas pequenos trechos derivados de elementos transponíveis ativos. Os transposons de DNA constituem aproximadamente 3% do genoma humano. Eles são estruturalmente similares aos elementos transponíveis de bactérias (Cap. 18).

Como é observado com outros elementos genéticos transponíveis (Cap. 18), a transposição de elementos Alu para genes e rearranjos de DNA causada por recombinação entre dois elementos Alu pode produzir fenótipos mutantes. Alelos mutantes que causam acolinesterasemia (uma deficiência da enzima colinesterase que resulta em paralisia transitória por exposição a relaxantes musculares) e neurofibromatose (tumores do tecido nervoso sob a pele e no sistema nervoso central) foram comprovados como resultantes da inserção de seqüências Alu em genes. Além disso, foi demonstrado que vários casos de hipercolesterolemia (quantidades excessivas de colesterol no sangue devido a uma deficiência de receptores de colesterol nas células do fígado) e de talasse-

mia (deficiências de hemoglobina) resultam de rearranjos gênicos produzidos por recombinação entre seqüências Alu proximamente ligadas. Estes e outros resultados mostraram claramente que as seqüências Alu têm um papel importante na evolução do genoma humano.

PONTOS IMPORTANTES

■ Os genomas eucarióticos contêm seqüências repetidas de DNA, com algumas seqüências presentes um milhão de vezes ou mais.

■ As seqüências de DNA em genomas eucarióticos estão comumente agrupadas em três classes: (1) seqüências únicas ou de cópia única presentes em uma ou poucas cópias por genoma, (2) seqüências moderadamente repetitivas presentes em 10 a 10^5 cópias e (3) seqüências altamente repetitivas presentes em mais de 10^5 cópias por genoma.

■ Duas famílias de elementos transponíveis moderadamente repetidos constituem 10% do genoma humano e desempenham um papel importante em sua evolução.

Exercícios Básicos Ilustram a análise genética básica.

1. Que diferenças nas estruturas químicas do DNA e proteínas permitem que os cientistas marquem uma ou outra destas macromoléculas com um isótopo radioativo?

Resposta: O DNA contém fósforo (o isótopo comum é ^{31}P) mas não enxofre; o DNA pode ser marcado cultivando-se células em meio contendo o isótopo radioativo do fósforo, ^{32}P . As proteínas contêm enxofre (o isótopo comum é ^{32}S) mas geralmente pouco ou nenhum fósforo. As proteínas podem ser marcadas cultivando-se células em meio contendo o isótopo radioativo do enxofre, ^{35}S .

2. Se a seqüência de um filamento de uma dupla hélice de DNA é ATCG, qual é a seqüência do outro filamento?

Resposta: Como os dois filamentos de uma dupla hélice são complementares – adenina sempre pareada com timina e guanina sempre pareada com citosina –, a seqüência do segundo filamento pode ser deduzida pela seqüência do primeiro filamento. Para ATCG, a dupla hélice terá a seguinte estrutura:

```

    ATCG
    ....
    TAGC
  
```

3. De que modo a seqüência do filamento complementar na dupla hélice do exercício 2 pode ser escrita como um único filamento de DNA?

Resposta: Lembre que os dois filamentos de uma dupla hélice de DNA têm polaridade química oposta; um filamento tem polaridade $5' \rightarrow 3'$ e o outro tem polaridade $3' \rightarrow 5'$, quando ambos são lidos na mesma direção. Como a convenção aceita é escrever as seqüências começando com a ponta $5'$ à esquerda e terminando com a ponta $3'$ à direita, o filamento de cima da dupla hélice deve ser escrito $5'-\text{ATCG}-3'$ e o filamento complementar, $5'-\text{CGAT}-3'$. A estrutura da dupla hélice deve ser escrita:

```

    5'-ATCG-3'
    .....
    3'-TAGC-5'
  
```

4. Se uma mistura de DNA e proteína é mostrada como possuidora de informação genética por algum ensaio tal como transformação em bactérias, como um pesquisador pode determinar se tal informação genética está presente no DNA ou no componente protéico?

Resposta: A especificidade biológica das enzimas fornece uma poderosa ferramenta para uso em muitas pesquisas. A enzima desoxirribonuclease (DNase) degrada DNA em mononucleotídeos e proteases degradam proteínas em componentes menores. Se a mistura de DNA e proteína é tratada com DNase e a informação genética é destruída, ela está estocada no DNA. Se a mistura é tratada com protease e a informação genética é perdida, ela reside no componente protéico da mistura.

5. Como um pesquisador determina onde uma seqüência específica de DNA está situada em um cromossomo?

Resposta: Em condições apropriadas, as seqüências complementares de DNA são pareadas para formar hélices duplas. Este processo é chamado helicoidização ou hibridização. Se uma seqüência específica de DNA é isolada e marcada com um isótopo radioativo ou um corante fluorescente, ela pode ser usada para localizar seqüências de DNA complementares nos cromossomos. O procedimento é chamado hibridização *in situ*. O DNA no cromossomo é desnaturado por exposição a alta temperatura ou pH e então deixado hibridizar com o filamento complementar de DNA ou RNA marcado (a "sonda"). Após ter ocorrido hibridização, o filamento marcado ou sonda no cromossomo é localizado por auto-radiografia ou microscopia de fluorescência (veja Enfoque Técnico: Hibridização *In Situ*).

Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. A alga vermelha *Polyides rotundus* estoca sua informação genética na dupla hélice de DNA. Quando o DNA foi extraído de células de *P. rotundus* e analisado, observou-se que 32% das bases eram

guaninas. Por esta informação, você pode determinar que porcentagem das bases neste DNA eram timinas? Caso sim, qual a porcentagem? Caso não, por que não?

Resposta: Os dois filamentos de uma dupla hélice de DNA são complementares um ao outro, com guanina (G) em um filamento sempre pareando com citosina (C) no outro, e, do mesmo modo, adenina (A) sempre pareada com timina (T). Portanto, as concentrações de G e C são sempre iguais, como o são as concentrações de A e T. Se 32% das bases na dupla hélice de DNA são G, então outros 32% são C. Juntos, G e C somam 64% das bases do DNA de *P. rotundus*. Logo, 36% das bases são A e T. Como a concentração de A deve ser igual à concentração de T, 18% ($36\% \times \frac{1}{2}$) das bases devem ser T.

2. O vírus ϕ X174 de *E. coli* estoca sua informação genética em DNA unifilar. Quando o DNA foi extraído do ϕ X174 e analisado, 21% das bases eram de unidades G. A partir desta informação, você pode determinar que porcentagem de bases neste DNA eram timinas? Caso sim, qual a porcentagem? Caso não, por quê?

Resposta: Não! A relação $A = T$ e $G = C$ só ocorre em moléculas bifilamentares de DNA devido a seus filamentos complementares. Como pareamento de bases não ocorre ou ocorre apenas como pareamento limitado intrafilamentar em ácidos

nucleicos unifilamentares, você não pode determinar a porcentagem de nenhuma das outras três bases pelo conteúdo de G do DNA de ϕ X 174.

3. Se cada cromossomo humano em estágio G_1 contém uma única molécula de DNA, quantas moléculas de DNA estariam presentes nos cromossomos do núcleo de (a) um ovócito humano, (b) um espermatozóide humano, (c) uma célula somática diplóide humana no estágio G_1 , (d) uma célula somática humana diplóide no estágio G_2 (e) um ovócito primário humano?

Resposta: Uma célula humana haplóide normal contém 23 cromossomos e uma célula humana diplóide normal contém 46 cromossomos, ou 23 pares de homólogos. Se os cromossomos antes da replicação contêm uma única molécula de DNA, os cromossomos após a replicação conterão duas moléculas de DNA, uma em cada uma das duas cromátides. Assim, ovócitos e espermatozóides humanos normais contêm 23 moléculas de DNA cromossômico. As células somáticas diplóides contêm 46 e 92 moléculas de DNA cromossômico nos estágios G_1 e G_2 , respectivamente; e um ovócito primário contém 92 de tais moléculas de DNA.

Questões e Problemas

Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 9.1. (a) Como os experimentos de transformação de Griffith diferem dos de Avery e seus colaboradores?
(b) Qual foi a contribuição significativa de cada um?
(c) Por que o trabalho de Griffith não é uma evidência do DNA como material genético, enquanto os experimentos de Avery e colaboradores deram uma prova direta de que o DNA leva a informação genética?
- 9.2. Um extrato livre de célula é preparado de células pneumocócicas tipo IIIS. Que efeito o tratamento desse extrato com (a) protease, (b) RNase e (c) DNase tem em sua capacidade subsequente de transformar células receptoras tipo IIR em tipo IIIS? Por quê?
- 9.3. Como pode ser demonstrado que misturar pneumococos tipo III mortos por aquecimento com tipo II vivos resulta em transferência de material genético do tipo III para o tipo II e não em restauração da viabilidade do tipo III pelo tipo II?
- 9.4. Qual a composição macromolecular de um vírus bacteriano ou bacteriófago tal como o fago T2?
- 9.5. (a) Qual o objetivo do experimento feito por Hershey e Chase?
(b) Como o objetivo foi atingido?
(c) Qual o significado desse experimento?
- 9.6. Como o experimento de reconstituição de Fraenkel-Conrat e colaboradores mostrou que a informação genética do vírus do mosaico do tabaco (TMV) é estocada em seu RNA e não em sua proteína?
- 9.7. (a) O que distingue um viróide de um vírus com RNA?
(b) O que distingue um príon de um viróide?
(c) De que modos virióides e príons são similares?
- 9.8. (a) De que material básico dispunham Watson e Crick para desenvolver um modelo de DNA?
(b) Qual sua contribuição para construir o modelo?
- 9.9. (a) Por que Watson e Crick escolheram uma dupla hélice para seu modelo de estrutura do DNA?
(b) Por que pontes de hidrogênio foram colocadas no modelo para conectar as bases?
- 9.10. (a) Se uma partícula de vírus contém DNA bifilar com 200.000 pares de bases, quantos nucleotídeos estariam presentes?
(b) Quantas espirais completas ocorreriam em cada filamento?
(c) Quantos átomos de fósforo estariam presentes?
(d) Qual seria o tamanho da configuração de DNA no vírus?
- 9.11. Se um filamento de DNA na dupla hélice de Watson-Crick tem uma sequência de bases de 5'-GTCATGAC-3', qual a sequência de bases do filamento complementar?
- 9.12. Quais as diferenças entre DNA e RNA?
- 9.13. DNA foi extraído de células de *Staphylococcus fermentans* e analisado quanto a composição de bases. Foi observado que 37% das bases são citosinas. Com esta informação, é possível prever que porcentagem das bases são adenina? Caso sim, qual porcentagem? Caso não, por que não?

- 9.14. RNA foi extraído do TMV (vírus do mosaico do tabaco) e encontrado conteúdo de 20% de citosina (20% das bases eram citosina). Com esta informação, é possível prever que porcentagem de bases no TMV são adeninas? Caso sim, que porcentagem? Caso não, por que não?
- 9.15. Indique se cada uma das seguintes citações sobre a estrutura do DNA é verdadeira ou falsa. (Cada letra é usada para se referir à concentração desta base no DNA.)
- $A + T = G + C$
 - $A = G$; $C = T$
 - $A/T = C/G$
 - $T/A = C/G$
 - $A + G = C + T$
 - $G/C = 1$
 - $A = T$ dentro de cada filamento.
 - As pontes de hidrogênio dão estabilidade à dupla hélice em citoplasmas aquosos.
 - As pontes hidrofóbicas dão estabilidade à dupla hélice em citoplasmas aquosos.
 - Quando separados, os dois filamentos de uma dupla hélice são idênticos.
 - Uma vez conhecida a sequência de bases de um filamento da dupla hélice de DNA, a sequência de bases do segundo filamento pode ser deduzida.
 - A estrutura de uma dupla hélice de DNA é invariante.
 - Cada par de nucleotídeos contém dois grupos fosfato, duas moléculas de desoxirribose e duas bases.
- 9.16. Os ácidos nucleicos de vários vírus foram extraídos e examinados para determinar sua composição de bases. Em vista dos seguintes resultados, o que você pode supor sobre a natureza física dos ácidos nucleicos desses vírus?
- 35% A, 35% T, 15% G e 15% C.
 - 35% A, 15% T, 25% G e 25% C.
 - 35% A, 30% U, 30% G e 5% C.
- 9.17. Compare e contraste as estruturas das formas A, B e Z de DNA.
- 9.18. A temperatura na qual metade de um filamento bifilamentar de DNA foi desnaturada é chamada temperatura de dissociação (*melting*), T_m . Por que a T_m depende diretamente do conteúdo de GC do DNA?
- 9.19. A relação entre a T_m e o conteúdo de GC pode ser expressa, em sua forma mais simplificada, pela fórmula $T_m = 69 + 0,41 (\% \text{ GC})$. (a) Calcule a temperatura de dissociação do DNA de *E. coli* que tem cerca de 50% de GC. (b) Avalie a % de GC do DNA de uma célula renal humana em que $T_m = 85^\circ\text{C}$.
- 9.20. As evidências disponíveis indicam que cada cromossomo eucariótico (excluindo cromossomos politênicos) contém uma única molécula gigante de DNA. Que níveis diferentes de organização desta molécula de DNA são aparentes em cromossomos de eucariontes em vários momentos durante o ciclo celular?
- 9.21. Uma planta diplóide de centeio, *Secale cereale*, tem $2n = 14$ cromossomos e aproximadamente $1,6 \times 10^{10}$ pb de DNA. O quanto de DNA está em um núcleo de uma célula de centeio em (a) metáfase mitótica, (b) metáfase I meiótica, (c) telófase mitótica e (d) telófase II meiótica?
- 9.22. Um núcleo diplóide de *Drosophila melanogaster* contém cerca de $3,4 \times 10^8$ pares de nucleotídeos. Suponha (1) que todo o DNA nuclear está acondicionado em nucleossomos e (2) que o tamanho médio do ligador internucleossômico é de 60 pares de nucleotídeos. Quantos nucleossomos estariam presentes em um núcleo diplóide de *D. melanogaster*? Quantas moléculas de histonas H2a, H2b, H3 e H4 seriam necessárias?
- 9.23. Os DNA satélites de *Drosophila virilis* podem ser isolados, essencialmente liberados da banda principal de DNA, por centrifugação em gradiente de densidade. Se esses DNA satélites forem fracionados em fragmentos de aproximadamente 40 nucleotídeos de tamanho e analisados em experimentos de desnaturação-renaturação, como você espera que sua cinética de hibridização seja comparada com a cinética de renaturação observada usando a banda principal de DNA similarmente fragmentada nas mesmas condições? Por quê?
- 9.24. Evidências experimentais indicam que a maioria das sequências de DNA altamente repetitivas nos cromossomos de eucariontes não geram nenhum RNA ou produtos protéicos. O que isto indica sobre a função do DNA altamente repetitivo?
- 9.25. (a) As (1) sequências de DNA de cópia única, (2) sequências de DNA moderadamente repetitivas, (3) sequências de DNA altamente repetitivas são mais prevalentes na eucromatina ou na heterocromatina?
- (b) A maioria das sequências de DNA de cópia única nos cromossomos são transcritas? As sequências de DNA altamente repetitivas na cromatina são transcritas?
- (c) Quais são as supostas funções da maioria (1) das sequências de DNA de cópia única, (2) das sequências de DNA altamente repetitivas?
- (d) Que papéis acredita-se que desempenham as sequências de DNA moderadamente repetitivas em eucariontes superiores?
- 9.26. Os cromossomos eucarióticos são metabolicamente mais ativos durante prófase, metáfase, anáfase, telófase ou interfase?
- 9.27. (a) Que funções têm (1) os centrômeros, (2) os telômeros?
- (b) Os telômeros têm características estruturais únicas?
- (c) Qual a função da telomerase?
- (d) Quando os cromossomos são quebrados por exposição a radiação de alta energia tal como os raios X, as pontas quebradas resultantes apresentam tendência pronunciada a se juntar e se fundir. Por que isto ocorre?
- 9.28. Que interesse especial tem a técnica biofísica de viscoelastometria para geneticistas?
- 9.29. Quantas moléculas de DNA estão presentes (a) nas regiões axiais, (b) nas alças laterais de cromossomos plumosos?

- 9.30. Os arcabouços de cromossomos eucarióticos são compostos de proteínas cromossômicas histônicas ou não histônicas? Como isto foi determinado experimentalmente?
- 9.31. (a) Que classe de proteínas cromossômicas, histonas ou não histonas, é a mais altamente conservada em diferentes espécies eucarióticas? Por que pode ser esperada esta diferença?
 (b) Se compararmos as proteínas cromossômicas histônicas e não histônicas de cromatina isolada de diferentes tecidos ou tipos de células de um determinado organismo eucariótico, que classe de proteínas apresentará a maior heterogeneidade? Por que não se espera que ambas as classes de proteínas sejam igualmente homogêneas em cromossomos de diferentes tecidos ou tipos de células?
- 9.32. A sequência do genoma humano revela que cerca de 44% do genoma consiste em elementos genéticos transponíveis ou sequências derivadas de elementos transponíveis.
 (a) Que tipos de elementos transponíveis estão presentes no genoma humano?
 (b) Todos os elementos transponíveis são ativos?
 (c) Por que os elementos Alu têm este nome?
 (d) Por que existem tantos elementos transponíveis no genoma humano?
- 9.33. Os elementos genéticos transponíveis são responsáveis por várias doenças humanas herdadas. Como um elemento transponível pode causar um distúrbio herdado?
- 9.34. (a) Se o genoma humano haplóide contém 3×10^9 pares de nucleotídeos e o peso molecular médio de um par de nucleotídeos é 660, quantas cópias do genoma humano estão presentes, em média, em 1 mg de DNA humano?
 (b) Qual o peso de uma cópia do genoma humano?
 (c) Se o genoma haplóide da pequena planta *Arabidopsis thaliana* contém $7,7 \times 10^7$ pares de nucleotídeos, quantas cópias do genoma de *A. thaliana* estão presentes, em média, em 1 mg de DNA de *A. thaliana*?
 (d) Qual o peso de uma cópia do genoma de *A. thaliana*?
 (e) Que importância têm os cálculos das alíneas anteriores para geneticistas?
- 9.35. Quando é analisado o DNA total de um rato canguru (*Dipodomys ordii*) por equilíbrio de centrifugação de gradiente de densidade, uma fração do DNA forma uma banda satélite distinta no gradiente. Este DNA satélite contém uma sequência altamente repetida de DNA que pode ser isolada do gradiente livre de outras sequências de DNA. Muitas dessas sequências altamente repetitivas de DNA estão localizadas em regiões de heterocromatina adjacentes aos centrômeros dos cromossomos ("heterocromatina centromérica"). Como um pesquisador poderia determinar onde a sequência satélite está localizada nos cromossomos desse rato?

Replicação do DNA e Cromossomos

Tópicos

- Características Básicas de Replicação do DNA *In Vivo*
- DNA-Polimerases e Síntese de DNA *In Vitro*
- O Complexo Aparelho de Replicação
- Aspectos Únicos de Replicação de Cromossomos Eucarióticos



Sete pares de gêmeos idênticos. Embora os gêmeos não tivessem recebido nenhuma instrução sobre como posar para a fotografia, note postura similar, posição das mãos e expressões faciais de ambos os membros de cada par.

Gêmeos Monozigóticos: Eles São Idênticos?

Desde o dia de seu nascimento, durante a infância, a adolescência e a vida adulta, Merry e Sherry foram confundidas uma com a outra. Quando estão separadas, Merry é chamada de Sherry em cerca de metade do tempo, e Sherry é confundida com Merry com igual frequência. Mesmo seus genitores têm dificuldade em distingui-las. Merry e Sherry são gêmeas monozigóticas ("idênticas"); ambas se desenvolveram de um único zigoto. Em um estágio inicial de clivagem, o embrião dividiu-se em duas massas celulares, e ambos os grupos de células se desenvolveram em embriões completos. Ambos os embriões se desenvolveram normalmente, e, em 7 de abril de 1955, uma foi denominada Merry e a outra Sherry.

As pessoas geralmente explicam os fenótipos quase idênticos de gêmeos monozigóticos como Merry e Sherry dizendo que elas "contêm os mesmos

genes". Claramente, isto não é verdade. Para ser exato, a frase deveria ser que os gêmeos idênticos contêm réplicas dos mesmos genes parentais. Mas este simples coloquialismo sugere que a maioria das pessoas, de fato, acreditam que as réplicas de um gene na prole realmente são idênticas. Se o genoma humano contém 25.000 a 35.000 genes, as réplicas de todos esses genes na prole são exatamente as mesmas em gêmeos idênticos?

A vida humana emerge de uma única célula, uma minúscula esfera com cerca de 0,1 mm de diâmetro. Esta célula dá origem a centenas de bilhões de células durante o desenvolvimento fetal. Na maturidade, um humano de tamanho médio contém cerca de 65 trilhões (65.000.000.000.000) de células. Com algumas exceções, cada uma dessas 65 trilhões de células contém uma réplica na prole de cada um dos 25.000 a 35.000 genes. Além disso, as células do corpo não são estáticas. Em alguns tecidos, células velhas estão continuamente sendo substituídas por células novas. Por exemplo, em uma pessoa saudável, as células da medula óssea produzem cerca de 2 milhões de hemácias por minuto. Embora nem todas as réplicas dos genes no corpo humano da prole sejam idênticas, o processo pelo qual tais genes são duplicados é muito preciso. O genoma humano haplóide contém cerca de 3×10^9 pares de nucleotídeos de DNA, todos os quais devem ser duplicados durante cada divisão celular. Neste capítulo, examinaremos como o DNA se replica e enfocaremos os mecanismos que garantem a fidelidade deste processo.

CARACTERÍSTICAS BÁSICAS DE REPLICAÇÃO DO DNA *IN VIVO*

A replicação do DNA ocorre semiconservativamente, é iniciada em origens únicas e em geral ocorre bidirecionalmente a partir de cada origem de replicação.



Em humanos, a síntese de um novo filamento de DNA ocorre a uma velocidade de cerca de 3.000 nucleotídeos por minuto. Em bactérias, cerca de 30.000 nucleotídeos são adicionados a uma cadeia de DNA nascente por minuto. Claramente, a maquinaria celular responsável pela replicação do DNA deve funcionar muito rápido, mas, e ainda mais importante, deve funcionar com grande precisão. De fato, a fidelidade de replicação do DNA é incrível, com uma média de apenas um erro por bilhão de nucleotídeos incorporados após a síntese e correção de erros durante e imediatamente após a replicação. Assim, a maioria dos genes de gêmeos idênticos são de fato idênticos, mas alguns terão mudado devido a erros de replicação e outros tipos de mutações (Cap. 13). A maioria das características principais do mecanismo pelo qual ocorre replicação rápida e precisa do DNA agora são conhecidas, embora muitos detalhes moleculares permaneçam por ser elucidados.

A síntese de DNA, como a síntese de RNA (Cap. 11) e proteínas (Cap. 12), envolve três etapas: (1) iniciação da cadeia, (2) ampliação, ou alongamento, da cadeia, e (3) término

da cadeia. Neste e nos dois capítulos seguintes, examinaremos os mecanismos pelos quais as células efetuam cada uma das três etapas na síntese destas importantes macromoléculas. Primeiro, entretanto, consideraremos algumas características importantes de replicação do DNA.

Replicação Semiconservativa

Quando Watson e Crick deduziram a estrutura da dupla hélice do DNA com seu pareamento complementar de bases, eles imediatamente reconheceram que a especificidade do pareamento de bases podia fornecer o princípio fundamental para um mecanismo simples de duplicação do DNA. Se os dois filamentos complementares de uma dupla hélice se separam, por rompimento das pontes de hidrogênio de cada par de bases, cada filamento parental poderia dirigir a síntese de um novo filamento complementar por causa das exigências de pareamento específico de bases (Fig. 10.1). Isto é, cada filamento parental poderia funcionar como um molde, um único filamento de DNA que especifica a sequência de nucleotídeos de um novo filamento complementar. A adenina, por exemplo, no filamento parental serviria como um molde mediante seu potencial de pontes de hidrogênio para a incorporação de timina no filamento complementar nascente. Este mecanismo de replicação do DNA é chamado **replicação semiconservativa** porque cada um dos filamentos complementares da dupla hélice parental é conservado (ou a dupla hélice é "semiconservada") durante o processo.

A ocorrência do mecanismo semiconservativo de Watson e Crick de replicação do DNA na bactéria *Escherichia coli* foi primeiro demonstrada por Matthew Meselson e Franklin

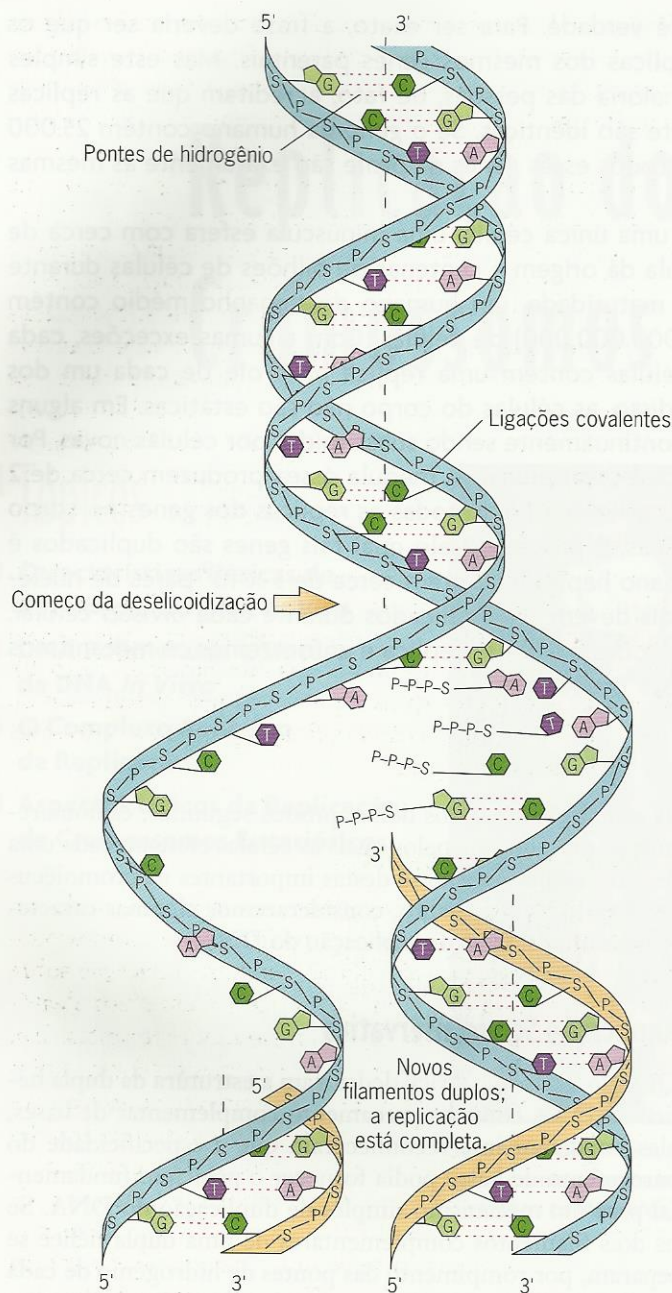


Fig. 10.1 ■ Modelo semiconservativo de replicação do DNA com base no pareamento complementar de bases proposto por Watson e Crick. Note que cada um dos filamentos parentais é conservado e serve como um molde para a síntese de um novo filamento complementar; isto é, a sequência de bases em cada filamento da prole é determinada pelos potenciais de pontes de hidrogênio das bases no filamento parental.

Stahl em 1958. Seu estudo definitivo é discutido em detalhes em *Um Marco na Genética: O DNA Replica-se Semiconservativamente*. Meselson e Stahl usaram o isótopo pesado de nitrogênio, ^{15}N , em lugar do nitrogênio leve normal, ^{14}N , para aumentar a densidade do DNA e então estudaram as mudanças na densidade durante replicação por centrifugação de gradiente de densidade. Veja em *Enfoque Técnico*:

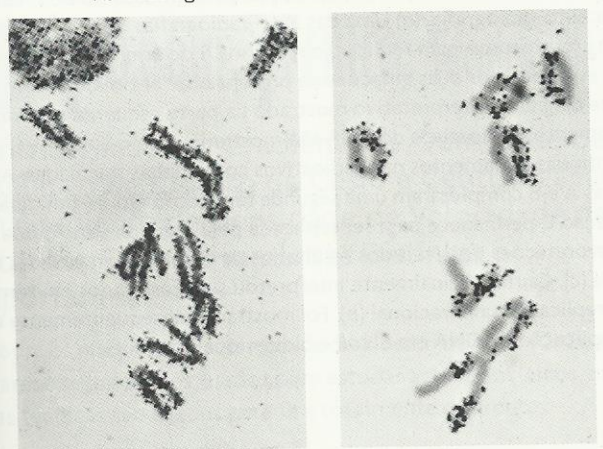
Técnicas de Centrifugação uma discussão sobre os tipos diferentes de gradientes de densidade e seu uso no isolamento de moléculas de DNA e no estudo de sua estrutura e replicação. Foi subsequentemente mostrado que a replicação do DNA ocorre semiconservativamente em vários outros microrganismos.

A replicação semiconservativa de cromossomos eucarióticos foi primeiro demonstrada em 1957 pelos resultados de experimentos feitos por J. Herbert Taylor, Philip Woods e Walter Hughes em células da ponta da raiz do feijão-fava, *Vicia faba*. Taylor e colaboradores marcaram os cromossomos de *Vicia faba* cultivando as pontas de raízes por oito horas (menos de uma geração celular) em meio contendo ^3H -timidina radioativa. As pontas da raiz foram então removidas do meio radioativo, lavadas e transferidas para um meio não radioativo contendo o alcalóide colchicina. É sabido que a colchicina liga-se a microtúbulos e impede a formação de fibras funcionais do fuso. Como resultado, os cromossomos filhos não sofrem sua separação anafásica normal. Assim, o número de cromossomos por núcleo duplica a cada ciclo celular na presença de colchicina. Essa duplicação do número de cromossomos a cada geração celular permitiu que Taylor e seus colaboradores determinassem quantas duplicações de DNA cada célula sofreu subsequentemente à incorporação de timidina radioativa. Na primeira metáfase em colchicina (c-metáfase), os núcleos conterão 12 pares de cromátides (ainda unidas pelo centrômero). Na segunda c-metáfase, os núcleos conterão 24 pares, e assim em diante.

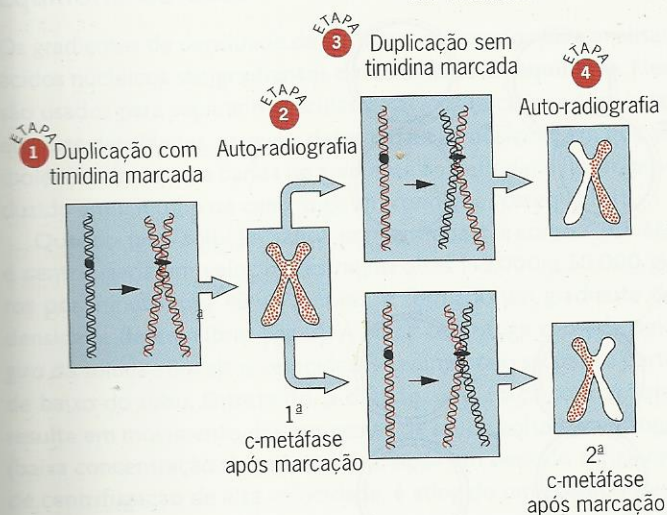
Quando a distribuição da radioatividade nos cromossomos de *Vicia faba* foi examinada por auto-radiografia (Cap. 9), ambas as cromátides de cada par foram similarmente marcadas na primeira c-metáfase (Fig. 10.2a). Entretanto, na segunda c-metáfase, apenas uma das cromátides de cada par era radioativa (Fig. 10.2b). Estes são exatamente os resultados esperados se a replicação do DNA for semiconservativa, havendo uma molécula de DNA por cromossomo (Fig. 10.2c). Em 1957, Taylor e seus colaboradores foram capazes de concluir que o DNA cromossômico em *Vicia faba* segregou de um modo semiconservativo durante cada divisão celular. A conclusão de que a dupla hélice se replicava semiconservativamente no feijão-fava teve que aguardar evidências subsequentes indicando que cada cromossomo contém uma única molécula de DNA. Experimentos análogos foram subsequentemente feitos com vários outros eucariotes, e, em todos os casos, os resultados indicaram que a replicação é semiconservativa.

Visualização das Forquilhas de Replicação por Auto-radiografia

A estrutura geral dos cromossomos bacterianos replicantes foi primeiro determinada por John Cairns em 1963, novamente por meio de auto-radiografia. Cairns cultivou bactérias *E. coli* em meio contendo ^3H -timidina por vários períodos de tempo, lisou as células de modo suave para não

Auto-radiografias de cromossomos de *Vicia faba*

(a) Primeira metáfase após replicação em ^3H -timidina. (b) Segunda metáfase após uma replicação adicional em ^3H -timidina.



(c) Interpretações das auto-radiografias acima em termos de replicação semiconservativa.

Fig. 10.2 ■ Resultados e interpretação do experimento de Taylor e colaboradores mostrando que o DNA cromossômico do feijão-fava, *Vicia faba*, se replica por um mecanismo semiconservativo.

quebrar os cromossomos (moléculas longas de DNA são sensíveis a agitação) e cuidadosamente coletou os cromossomos em membranas de filtragem. Esses filtros foram afixados a lâminas de vidro, revestidos com uma emulsão sensível a partículas β (os elétrons de baixa energia emitidos durante o decaimento do trítio) e guardados no escuro por um período de tempo para permitir decaimento radioativo suficiente. Quando os filmes foram revelados, as auto-radiografias (Fig. 10.3a) mostraram que os cromossomos de *E. coli* são estruturas circulares que existem como intermediários em forma de θ durante a replicação. As auto-radiografias indicaram ainda que a deselicoidização dos dois filamentos parentais complementares (que é necessária para sua separação) e sua replicação semiconservativa ocorrem simultaneamente ou são bem acopladas. Como a dupla hélice parental deve girar

360° para desenrolar cada giro da hélice, deve existir algum tipo de “rotação”. Geneticistas hoje sabem que a rotação necessária é uma quebra unifilar transitória (clivagem de uma ligação fosfodiéster em um filamento da dupla hélice) produzida pela ação de enzimas chamadas topoisomerases.

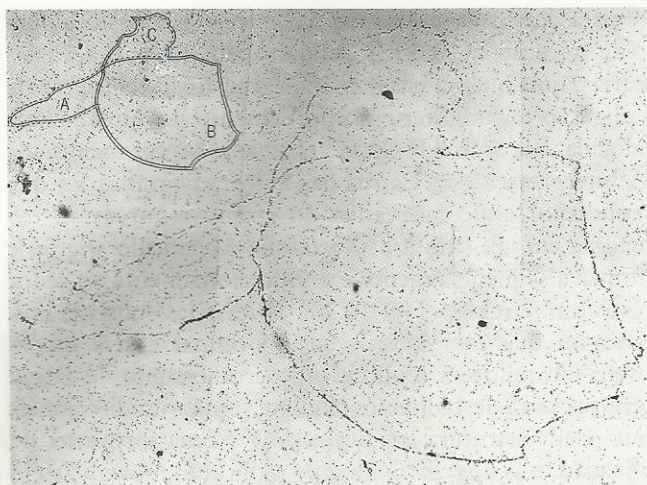
A interpretação de Cairns das auto-radiografias foi que a replicação semiconservativa do cromossomo de *E. coli* começava em um sítio específico, que é chamado *origem*, e continuava sequencialmente e unidirecionalmente ao redor da estrutura circular (Fig. 10.3b). Evidências subsequentes mostraram que esta interpretação original estava incorreta em um ponto: a replicação do cromossomo de *E. coli* ocorre na verdade bidirecionalmente, e não unidirecionalmente. Cada estrutura em forma de Y é uma **forquilha de replicação**, e as duas forquilhas de replicação movem-se em sentidos opostos sequencialmente ao redor do cromossomo circular (Fig. 10.3c).

A replicação bidirecional do cromossomo circular de *E. coli* discutida ocorre durante a divisão celular. Não deve ser confundida com a replicação de círculo rolante, que medeia a transferência de cromossomos das Hfr para as F^- (Cap. 8). Alguns cromossomos virais replicam-se pelo mecanismo de círculo rolante; veja a seção Replicação por Círculo Rolante, mais adiante neste capítulo.

Origens Únicas de Replicação

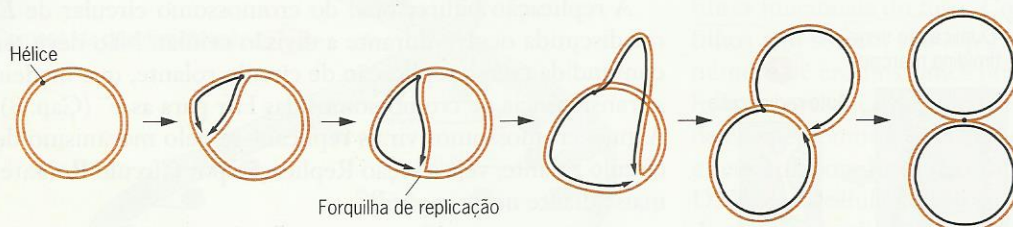
Os resultados de Cairns estabeleceram a existência de um sítio de iniciação ou **origem** de replicação no cromossomo circular de *E. coli*, mas não deram nenhuma indicação sobre se a origem era um sítio único ou ocorria em sítios aleatórios em uma população de cromossomos replicantes. Em cromossomos bacterianos e virais, geralmente há uma única origem por cromossomo, e essa única origem controla a replicação de todo o cromossomo. Nos cromossomos grandes de eucariontes, várias origens controlam coletivamente a replicação da molécula gigante de DNA presente em cada cromossomo. Evidências atuais indicam que essas várias origens de replicação em cromossomos eucarióticos também ocorrem em locais específicos. Cada origem controla a replicação de uma unidade de DNA chamada *répicon*. Assim, a maioria dos cromossomos procarióticos contém um único répicon, enquanto os cromossomos eucarióticos geralmente contêm muitos répicons.

A origem única de replicação, chamada *oriC*, no cromossomo de *E. coli* foi caracterizada em consideráveis detalhes. *OriC* tem 245 pares de nucleotídeos de tamanho e contém duas seqüências repetidas conservadas diferentes (Fig. 10.4). Uma seqüência com 13 pb está presente como três repetições em tandem. Estas três repetições são ricas em pares de bases A-T, facilitando a formação de uma região localizada de separação de filamento chamada **bolha de replicação**. Lembre que os pares de bases A-T são mantidos juntos por apenas duas pontes de hidrogênio, enquanto os pares de bases G-C têm três (Cap. 9). Assim, os dois filamentos de

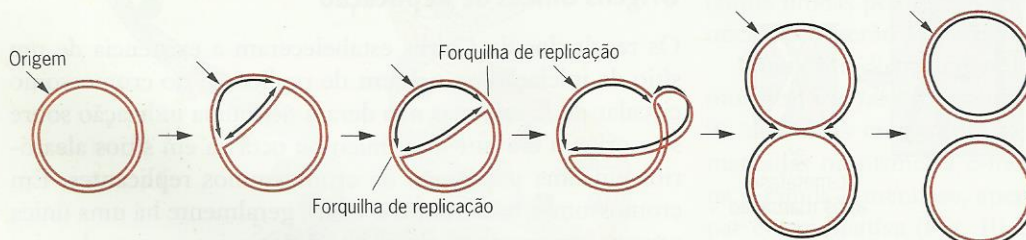


(a)

Fig. 10.3 ■ Visualização da replicação do cromossomo de *E. coli* por auto-radiografia. (a) Uma das auto-radiografias de Cairns de um cromossomo com replicação na forma θ de uma célula que foi cultivada por duas gerações na presença de ^3H -timidina, com este diagrama interpretativo mostrado na parte esquerda superior. Filamentos radioativos do DNA são mostrados como linhas contínuas e filamentos não radioativos como linhas tracejadas. As alças A e B completaram uma segunda replicação em ^3H -timidina; a seção C permanece para ser replicada pela segunda vez. As duas interpretações possíveis dos resultados de Cairns são mostradas em (b) e (c). Cairns originalmente interpretou seus resultados em termos de replicação unidirecional (b). Foi mostrado subsequentemente que a replicação do DNA em *E. coli* é bidirecional (c).



(b) Interpretação original: replicação unidirecional.



(c) Interpretação correta: replicação bidirecional.

regiões ricas em A-T do DNA separam-se mais facilmente, isto é, com menor gasto de energia. A formação de uma zona localizada de desnaturação é uma primeira etapa essencial na replicação de todos os DNA bifilamentares. Outro componente conservado de *oriC* é uma sequência de 9 pb que é repetida quatro vezes e é intercalada com outras sequências. Estas quatro sequências são sítios de ligação para uma proteína que tem um papel importante na formação da bolha de replicação. Mais adiante neste capítulo discutiremos detalhes adicionais do processo de iniciação da síntese de DNA nas origens e as proteínas que estão envolvidas.

As múltiplas origens de replicação em cromossomos eucarióticos também parecem ser sequências específicas de DNA. Na levedura *Saccharomyces cerevisiae*, segmentos de DNA cromossômico que permitem a um fragmento de DNA circular replicar-se como uma unidade independente (autonomamente), isto é, como uma unidade extracromossômica auto-replicante, foram identificados e caracterizados. Tais sequências são cha-

madas **elementos ARS** (para sequências de replicação autônoma). Sua frequência no genoma de levedura corresponde bem ao número de origens de replicação, e foi mostrado experimentalmente que algumas funcionam como origens. Os elementos ARS têm cerca de 50 pares de bases de tamanho e incluem um cerne de sequência de 11 pb rica em A-T,

ATTTATPuTTTA
TAAATAPiAAAT

(onde Pu é uma das duas purinas e Pi é uma das duas pirimidinas), e cópias imperfeitas adicionais desta sequência. A capacidade dos elementos ARS de funcionar como origens de replicação é abolida por mudanças de pares de bases dentro dessa sequência cerne conservada.

As tentativas de caracterizar as origens de replicação em eucariontes multicelulares foram amplamente infrutíferas. A despeito das evidências de que a replicação é iniciada em se-

ENFOQUE TÉCNICO: Técnicas de Centrifugação

Centrifugação em Gradiente de Densidade de Equilíbrio de Cloreto de Césio e Centrifugação em Gradiente de Densidade de Velocidade de Sacarose: Duas Ferramentas Distintas e Importantes

Embora os gradientes de densidade de cloreto de césio (CsCl) e de sacarose sejam ambos usados para estudar ácidos nucleicos, estas duas técnicas são usadas para fins totalmente diferentes.

Centrifugação em Gradiente de Densidade de Equilíbrio de CsCl

Os gradientes de densidade de CsCl que são usados para analisar ácidos nucleicos são **gradientes de densidade de equilíbrio**. Eles são usados para separar moléculas de ácidos nucleicos com base em suas densidades, as quais determinam as posições em que as moléculas farão uma banda no gradiente de densidade linear produzido em um tubo de centrifuga em condições de equilíbrio.

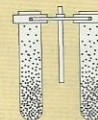
Quando uma solução salina concentrada tal como CsCl 6M é centrifugada em velocidades muito altas (30.000 a 50.000 giros por minuto) por 48 a 72 horas, é formado um gradiente de densidade de equilíbrio (Fig. 1). A força centrífuga causada pelo giro da solução em altas velocidades sedimenta o sal para a parte de baixo do tubo. Entretanto, a difusão, uma força contraposta, resulta em movimento das moléculas de sal de volta para o topo (baixa concentração de sal) do tubo. Após um período suficiente de centrifugação de alta velocidade, é atingido um equilíbrio entre sedimentação e difusão, quando existe um gradiente linear de densidade crescente de cima para baixo no tubo.

A densidade do CsCl 6M é de cerca de $1,7 \text{ g/cm}^3$; em equilíbrio, o gradiente de densidade linear produzido por centrifugação em alta velocidade de uma solução de CsCl 6M irá variar de cerca de $1,65 \text{ g/cm}^3$ em cima do tubo até cerca de $1,75 \text{ g/cm}^3$ na parte de baixo do tubo. As densidades dos ácidos nucleicos de ocorrência mais natural caem dentro desta faixa. Se DNA está presente em tal gradiente, ele irá mover-se para uma posição onde a densidade da solução salina é igual a sua própria densidade. Assim, se uma mistura de DNA de *E. coli* contendo o isótopo pesado de nitrogênio, ^{15}N , e DNA de *E. coli* contendo o isótopo leve normal de nitrogênio, ^{14}N , for submetida a centrifugação de gradiente de densidade de equilíbrio de CsCl, as moléculas de DNA irão separar-se em duas "bandas", uma consistindo em DNA "pesado" (contendo ^{15}N) e a outra em DNA "leve" (contendo ^{14}N).

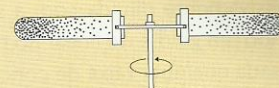
Centrifugação em Gradiente de Densidade de Velocidade de Sacarose

Os gradientes de densidade de sacarose que são usados para analisar ácidos nucleicos são **gradientes de densidade de velocidade**.

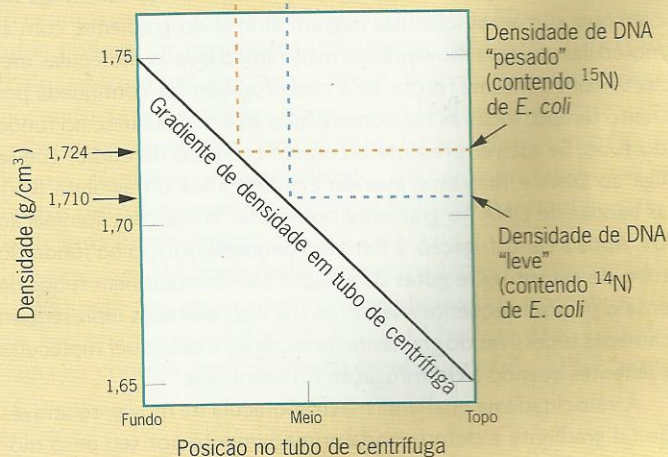
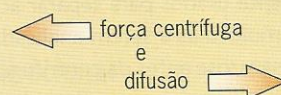
ETAPA 1 Preparar solução de CsCl 6M e adicionar mistura de DNA contendo ^{14}N e ^{15}N .



ETAPA 2 Centrifugar a 50.000 rpm por 48 a 72 horas.



É estabelecido um equilíbrio entre



ETAPA 3 Fazer furo no fundo do tubo de centrifuga e coletar frações.

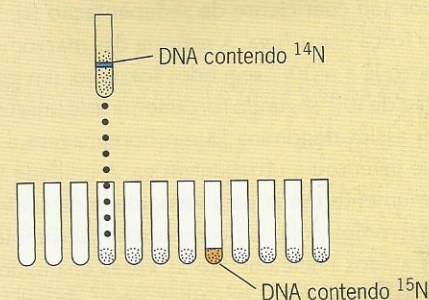


Fig. 1 ■ Gradiente de densidade de equilíbrio de CsCl.



ENFOQUE TÉCNICO: Técnicas de Centrifugação (continuação)

Eles são usados para avaliar os tamanhos de moléculas de ácidos nucleicos com base em suas taxas ou velocidades de movimento através de um gradiente de densidade de sacarose pré-formado quando exposto a uma força centrífuga.

Um marcador de gradiente é usado para preparar um gradiente de densidade pré-formado em um tubo de centrifuga (Fig. 2). O marcador de gradiente contém duas câmaras para soluções de sacarose. Uma câmara é cheia com uma solução de sacarose da concentração desejada no fundo do tubo da centrifuga; a outra é cheia com uma solução de sacarose com a concentração desejada no topo do tubo. À medida que o tubo é enchido, a solução de sacarose em alta concentração é progressivamente diluída com a solução de baixa concentração, produzindo um gradiente linear no tubo.

Embora possa ser usada qualquer faixa de concentrações de sacarose, 5 a 20 por cento e 10 a 40 por cento dos gradientes de sacarose são amplamente usados para separar ácidos nucleicos. A solução contendo as moléculas a serem analisadas é colocada no topo do gradiente de sacarose antes da centrifugação. A densidade desta solução deve ser mais baixa que a densidade de sacarose no topo do tubo de centrifuga de modo que ela ficará na superfície do gradiente de sacarose. O tubo é então colocado em um suporte do rotor giratório e centrifugado em alta velocidade (20.000 a 30.000 giros por minuto) por uma a algumas horas. A força centrífuga faz com que as macromoléculas migrem através do gradiente, com as moléculas maiores movendo-se mais rápido que as moléculas menores com a mesma forma. Se a centrifugação for continuada por muito tempo, todas as macromoléculas irão acumular-se no fundo do tubo. Se estiver presente o corante brometo de etídio, que se liga ao DNA e fluoresce quando exposto a luz ultravioleta (UV), as bandas de DNA no gradiente podem ser vistas com iluminação UV. Após a centrifugação, é feito um pequeno orifício no fundo do tubo de centrifuga, e gotas da solução são coletadas em tubos de ensaio para subsequente análise. As frações coletadas representam camadas sucessivas do gradiente de sacarose e quaisquer moléculas existentes quando a centrifugação foi terminada.

A velocidade na qual uma macromolécula irá mover-se através de tal gradiente é determinada primariamente por seu peso molecular e forma. A velocidade de movimento das moléculas através de uma solução de menor densidade durante sedimentação é medida por seu coeficiente de sedimentação (s), que é igual a velocidade/força centrífuga. A maioria das macromoléculas tem valores s entre 10^{-13} e 10^{-11} segundo. Portanto, um coeficiente de sedimentação de 10^{-13} segundo foi denominado uma unidade Svedberg (S). Foi assim denominado em homenagem a Svedberg, o cientista que inventou a ultracentrifuga. A unidade Svedberg é a unidade comumente usada para velocidade de sedimentação ao lidar com macromoléculas ou complexos macromoleculares. Por exemplo, os ácidos nucleicos são às vezes chamados moléculas 5S ou 10S, e as subunidades dos ribossomos de procariontes são freqüentemente referidas como subunidades 30S e 50S.

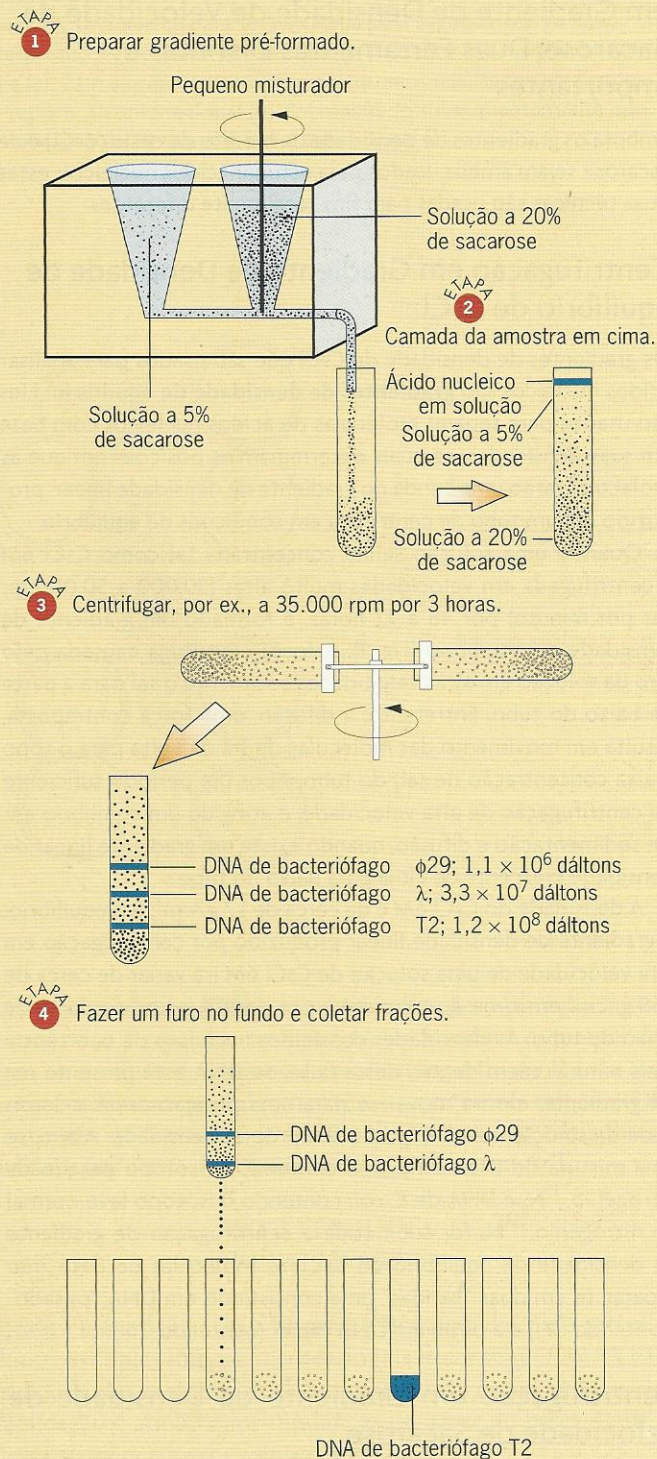


Fig. 2 ■ Gradiente de densidade de velocidade de sacarose.

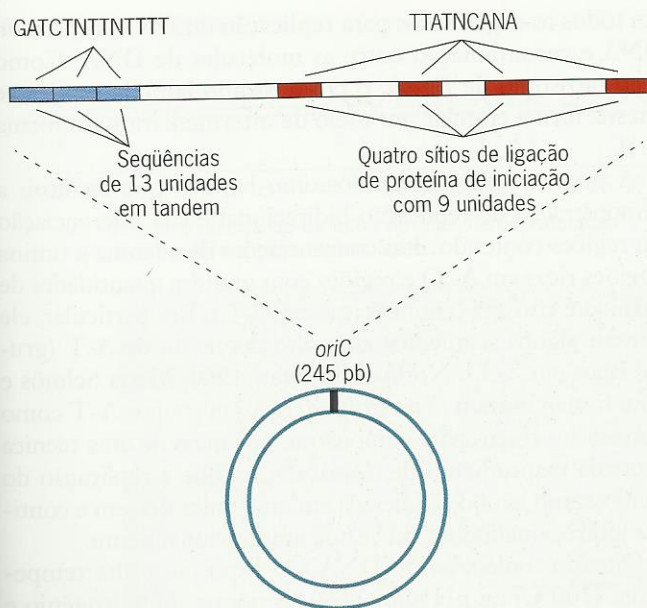


Fig. 10.4 ■ Estrutura de *oriC*, a única origem de replicação no cromossomo de *E. coli*.

seqüências específicas *in vivo* e da disponibilidade das seqüências de genomas inteiros, os componentes de uma origem funcional ainda não estão claros. Parece haver dois motivos principais para essa falha em identificar as origens de replicação. Primeiro, os testes funcionais usados em leveduras – a capacidade da origem de apoiar a replicação de um plasmídeo ou cromossomo artificial – não deram resultados confiáveis em outros eucariontes. As seqüências que apóiam a replicação de plasmídeos em células de mamíferos, por exemplo, em geral resultam no início de replicação em sítios aleatórios ou múltiplos. Segundo, evidências consideráveis sugerem hoje que o início da replicação envolve seqüências relativamente longas de DNA, de até milhares de pares de bases, em metazoários, dificultando a caracterização das origens.

Um enfoque bem-sucedido para caracterizar as origens de replicação que funcionam em eucariontes superiores foi

estudar a replicação de vírus com DNA. A replicação do vírus 40 de símios (SV40) foi amplamente estudada. O SV40 reproduz-se em células de primatas que crescem em cultura. O genoma do SV40 é uma dupla hélice circular de DNA contendo apenas 5.243 pares de nucleotídeos, o que o torna ideal para estudos de replicação do DNA em eucariontes. A maior parte dos detalhes de replicação do SV40 são conhecidos. Uma seqüência de 64 pares de nucleotídeos, o cerne da origem (Fig. 10.5), é tudo o que é necessário para iniciar replicação do DNA em células infectadas. O cerne de origem contém um elemento central com 27 pares de nucleotídeos flanqueados por uma seqüência rica em AT e outra seqüência. A replicação é iniciada quando uma proteína viral chamada grande antígeno T se liga à região do cerne de origem e, com a ajuda de proteínas celulares, inicia a deselicoi-dização dos filamentos de DNA na origem. Curiosamente, alguns vírus com DNA grande tais como o Epstein-Barr contêm três origens de replicação. Uma origem é usada para replicação durante infecções latentes e as outras duas origens controlam a replicação em infecções produtivas, ou líticas.

Replicação Bidirecional

Cairns interpretou seus auto-radiogramas de replicação dos cromossomos de *E. coli* com base na visão prevalente de que a replicação começa em uma única origem e que uma única forquilha de replicação se move unidirecionalmente ao redor do cromossomo circular (Fig. 10.3b). Entretanto, seus resultados podem ser explicados igualmente bem pela formação de duas forquilhas na origem única, com as duas forquilhas movendo-se em direções opostas, ou bidirecionalmente, ao redor do cromossomo circular (Fig. 10.3c). Hoje, temos evidências definitivas mostrando que a replicação em *E. coli* e muitos outros procariontes ocorre bidirecionalmente a partir de uma única origem.

A replicação bidirecional foi primeiro demonstrada convincentemente por experimentos com alguns dos pequenos vírus bacterianos que infectam *E. coli*. O bacteriófago lambda (fago λ) contém uma única molécula linear de DNA com

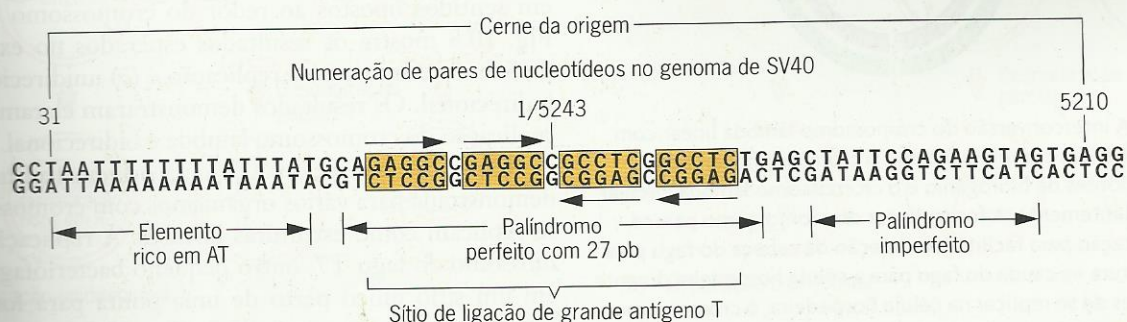


Fig. 10.5 ■ O cerne da origem de replicação do vírus 40 de símios (SV40). O cerne da origem é a seqüência mínima que pode iniciar replicação *in vivo*. Palíndromos são seqüências de DNA com a mesma seqüência de nucleotídeos quando lida em sentido oposto a partir de um ponto central de simetria. Eles são geralmente os sítios de ligação para proteínas que interagem com o DNA, tais como o grande antígeno T do SV40, que é a primeira proteína a se ligar ao cerne de origem durante o início da replicação do SV40.

apenas 17,5 μm de tamanho. O cromossomo do fago λ é um tanto incomum porque tem uma região unifilamentar, com 12 nucleotídeos de tamanho, na ponta 5' de cada filamento complementar (Fig. 10.6). Essas pontas unifilamentares, chamadas “coesivas” ou “adesivas”, são complementares umas às outras. As pontas coesivas de um cromossomo lambda podem, portanto, formar pares de bases em uma estrutura circular com pontes de hidrogênio. Um dos primeiros eventos a ocorrer após um cromossomo lambda ser injetado em uma célula hospedeira é a sua conversão em uma molécula circular covalentemente fechada (Fig. 10.6). Essa conversão da forma circular com pontes de hidrogênio para a forma circular fechada covalentemente é catalisada pela *DNA-ligase*, uma enzima importante que fecha quebras unifilamentares em duplas hélices de DNA. A *DNA-ligase* é necessária

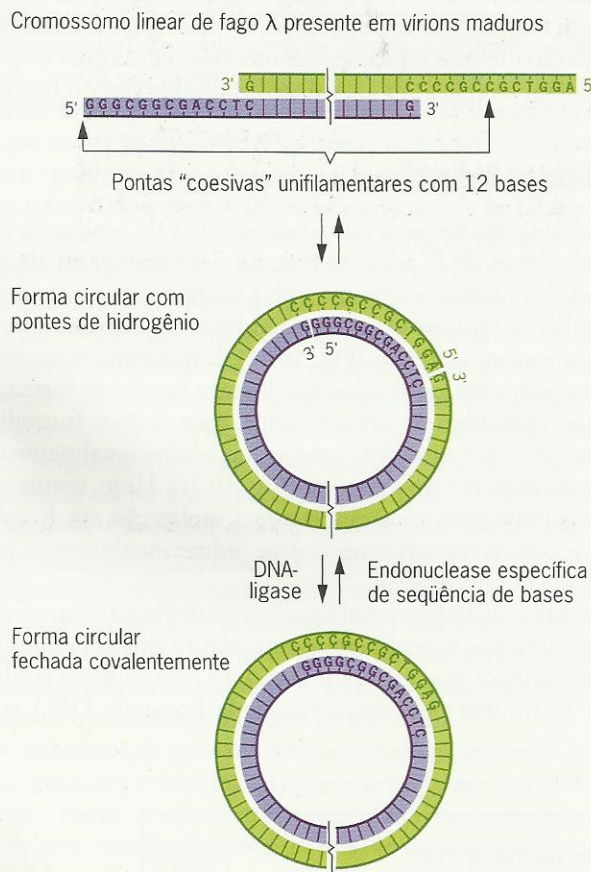


Fig. 10.6 ■ A interconversão do cromossomo lambda linear com suas pontas coesivas complementares, o cromossomo lambda circular com pontes de hidrogênio e o cromossomo lambda circular fechado covalentemente. A forma linear do cromossomo parece ser uma adaptação para facilitar sua injeção da cabeça do fago pela pequena abertura na cauda do fago para a célula hospedeira durante infecção. Antes de se replicar na célula hospedeira, o cromossomo é convertido para a forma circular fechada covalentemente. São mostradas apenas as pontas do cromossomo do fago maduro. A linha vertical com uma ondulação indica que a parte central do cromossomo não é mostrada. Todo cromossomo lambda tem cerca de 48.502 pares de nucleotídeos de tamanho.

em todos os organismos para replicação do DNA, reparo do DNA e recombinação entre as moléculas de DNA. Como o cromossomo de *E. coli*, o cromossomo lambda replica-se em sua forma circular por meio de intermediários em forma de θ .

A característica do cromossomo lambda que facilitou a demonstração de replicação bidirecional é sua diferenciação em regiões contendo altas concentrações de adenina e timina (regiões ricas em A-T) e regiões com grandes quantidades de guanina e citosina (regiões ricas em G-C). Em particular, ele contém alguns segmentos com alto conteúdo de A-T (grupos ricos em A-T). No final dos anos 1960, Maria Schnös e Ross Inman usaram esses grupos ricos em grupos A-T como marcadores físicos para demonstrar, por meio de uma técnica chamada mapeamento de desnaturação, que a replicação do cromossomo lambda é iniciada em uma única origem e continua bidirecionalmente em vez de unidirecionalmente.

Quando moléculas de DNA são expostas a alta temperatura (100°C) ou pH alto (11,4), as pontes de hidrogênio e ligações hidrofóbicas que unem os filamentos complementares na configuração de dupla hélice são quebradas, e os dois filamentos separam-se, um processo chamado desnaturação. Como os pares de bases A-T são mantidos juntos por apenas duas pontes de hidrogênio, em comparação com três pontes de hidrogênio nos pares G-C, moléculas ricas em A-T desnaturam-se mais facilmente (em pH ou temperaturas mais baixas) que moléculas ricas em G-C. Quando cromossomos lambda são expostos a pH de 11,05 por 10 minutos em condições apropriadas, os grupos ricos em A-T desnaturam-se para formar bolhas de desnaturação, que são detectáveis por microscopia eletrônica, enquanto as regiões ricas em G-C permanecem no estado de dupla hélice (Fig. 10.7). Tais bolhas de desnaturação podem ser usadas como marcadores físicos, esteja o cromossomo lambda em sua forma linear madura, em sua forma circular ou em seus intermediários replicativos em forma de θ . Examinando as posições dos pontos de ramificação (estruturas em forma de Y) relativas às posições das bolhas de desnaturação em um número maior de intermediários de replicação em forma de θ , Schnös e Inman demonstraram que ambos os pontos de ramificação são forquilha de replicação que se movem em sentidos opostos ao redor do cromossomo circular. A Fig. 10.8 mostra os resultados esperados no experimento de Schnös e Inman se a replicação é (a) unidirecional ou (b) bidirecional. Os resultados demonstraram claramente que a replicação do cromossomo lambda é bidirecional.

Replicação bidirecional de uma origem fixa também foi demonstrada para vários organismos com cromossomos que se replicam como estruturas lineares. A replicação do cromossomo do fago T7, outro pequeno bacteriófago, começa em um sítio único perto de uma ponta para formar uma estrutura em forma de “olho” (Fig. 10.9a) e então continua bidirecionalmente até que uma forquilha atinja a ponta mais próxima. A replicação da estrutura em forma de Y (Fig. 10.9b) continua até que a segunda forquilha atinja a outra ponta da molécula, produzindo dois cromossomos da prole.

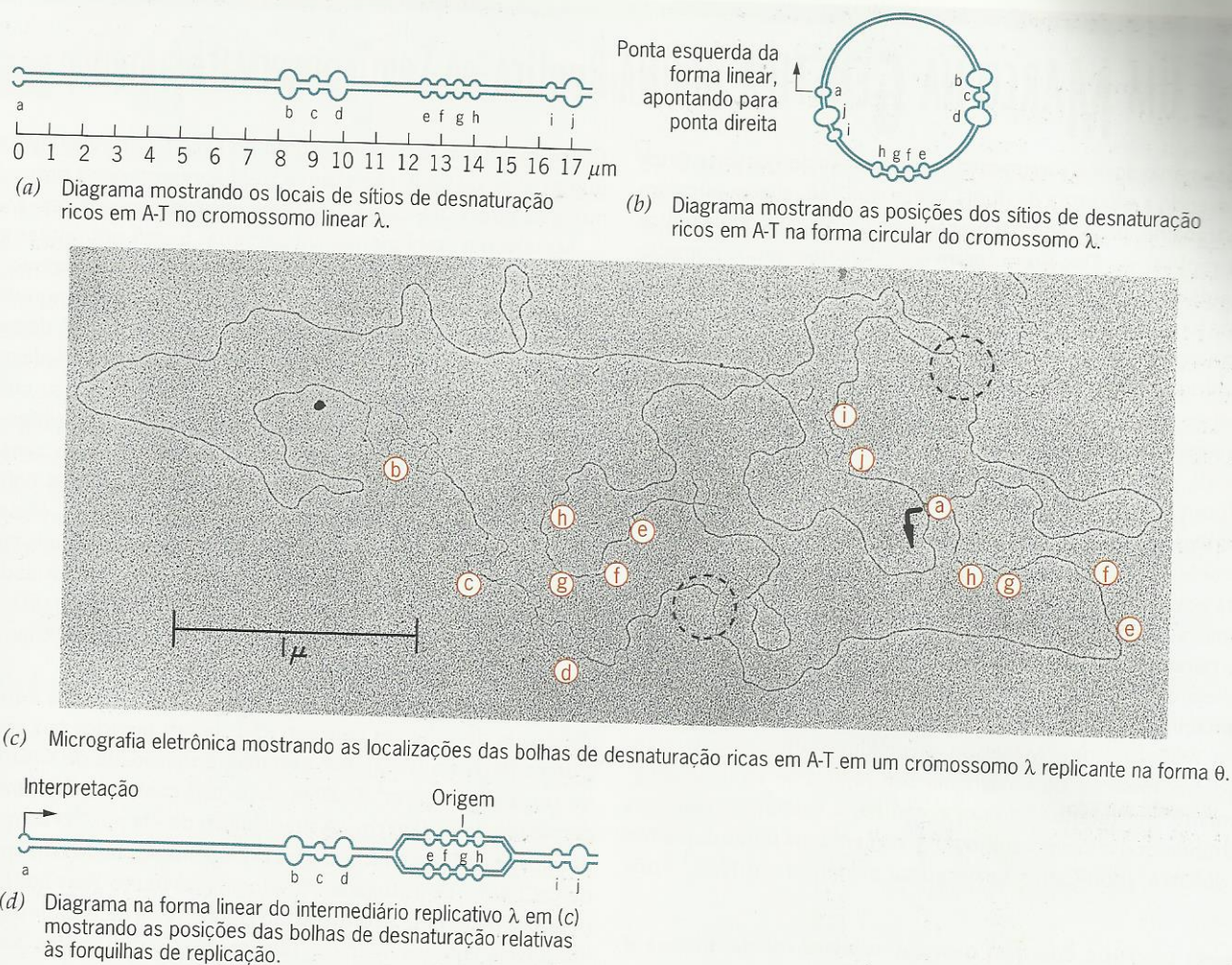


Fig. 10.7 ■ Ilustração do uso dos sítios de desnaturação ricos em A-T como marcadores físicos para provar que o cromossomo do fago λ se replica bidirecionalmente, e não unidirecionalmente.

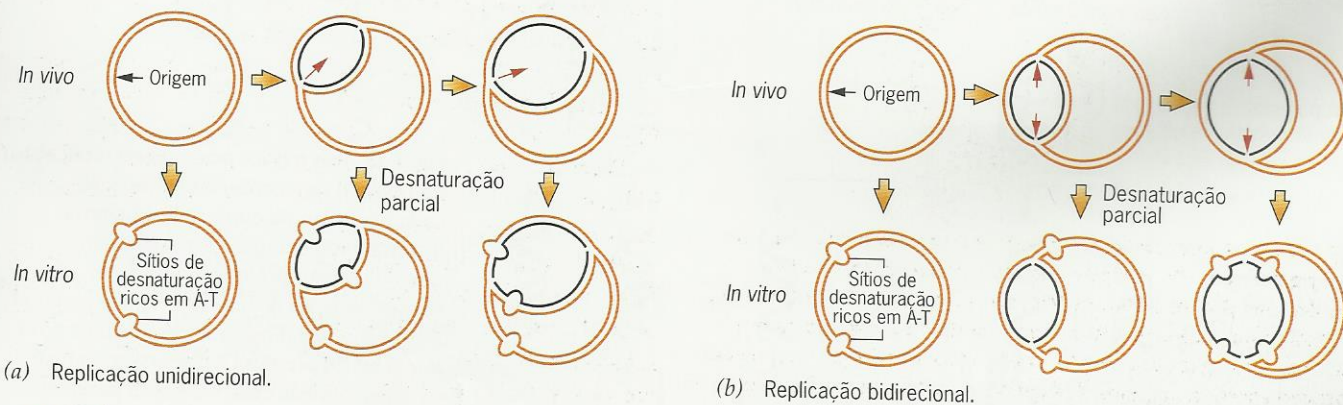


Fig. 10.8 ■ Lógica do procedimento de mapeamento de desnaturação usado por Schnös e Inman para distinguir os modos (a) unidirecional e (b) bidirecional de replicação cromossômica.



UM MARCO NA GENÉTICA: O DNA Replica-se Semiconservativamente

Cinco semanas após o surgimento da publicação de Watson e Crick descrevendo a estrutura da dupla hélice do DNA, eles publicaram outro trabalho¹ propondo um mecanismo pelo qual a dupla hélice podia replicar-se. Eles propuseram que dois filamentos complementares da dupla hélice desenrolam-se e se separam e que cada filamento orienta a síntese de um novo filamento complementar. A sequência de bases em cada filamento parental é usada como um molde, e as restrições de pareamento de bases dentro da dupla hélice – adenina com timina e guanina com citosina – ditam a sequência de bases no filamento recém-sintetizado. O mecanismo de Watson-Crick hoje é chamado *replicação semiconservativa* (porque a molécula parental é metade conservada) para distingui-lo de outros possíveis mecanismos de replicação (Fig. 1). Na replicação conservativa, a dupla hélice parental seria conservada, e uma nova prole de dupla hélice seria sintetizada. Na replicação dispersiva, segmentos de ambos os filamentos da molécula de DNA parental seriam conservados e usados como moldes para a síntese de segmentos complementares que subsequente seriam unidos para produzir filamentos de DNA da prole.

Em 1958, Matthew Meselson e Franklin Stahl demonstraram que o cromossomo de *Escherichia coli* replica-se semiconservativamente.² Entretanto, na época, não havia evidências decisivas mostrando que o cromossomo de *E. coli* era uma única dupla hélice de DNA. John Cairns forneceu tal evidência em 1962.³ Após

Cairns mostrar que o cromossomo de *E. coli* continha uma única dupla hélice de DNA, Meselson e Stahl apresentaram a prova definitiva de que o DNA se replica semiconservativamente em *E. coli*.

Meselson e Stahl cultivaram bactérias *E. coli* por muitas gerações em um meio no qual o isótopo pesado de nitrogênio, ¹⁵N, foi substituído pelo isótopo normal leve, ¹⁴N. As bases purínicas e pirimidínicas no DNA contêm nitrogênio. Assim, o DNA de bactérias cultivadas em meio contendo ¹⁵N terão uma densidade maior (massa por unidade de volume) do que o DNA de células cultivadas em meio contendo ¹⁴N. Moléculas com densidades diferentes podem ser separadas por um procedimento chamado **centrifugação de gradiente de densidade de equilíbrio** (veja Enfoque Técnico: Técnicas de Centrifugação). Como resultado, Meselson e Stahl foram capazes de distinguir entre os três possíveis modos de replicação do DNA seguindo as mudanças na densidade do DNA de células cultivadas em meio com ¹⁵N e então transferidas para meio com ¹⁴N por vários períodos de tempo, os chamados experimentos de transferência de densidade.

A densidade da maioria dos DNA é mais ou menos a mesma densidade de soluções concentradas de sais pesados tais como o cloreto de cério (CsCl). Por exemplo, a densidade de CsCl 6M é de cerca de 1,7 g/cm³. O DNA de *E. coli* contendo ¹⁴N tem uma densidade de 1,710 g/cm³. A substituição de ¹⁵N por ¹⁴N aumenta a densidade do DNA de *E. coli* para 1,724 g/cm³. Quando uma solução de CsCl 6M é centrifugada em velocidades muito altas por longos períodos de tempo, é formado um gradiente de densidade de equilíbrio (veja Enfoque Técnico: Técnicas de Centrifugação). Se DNA está presente em tal gradiente, ele irá mover-se para uma posição em que a densidade de CsCl é igual a sua própria densidade.

Meselson e Stahl usaram bactérias que haviam sido cultivadas em meio contendo ¹⁵N por várias gerações (e portanto continham

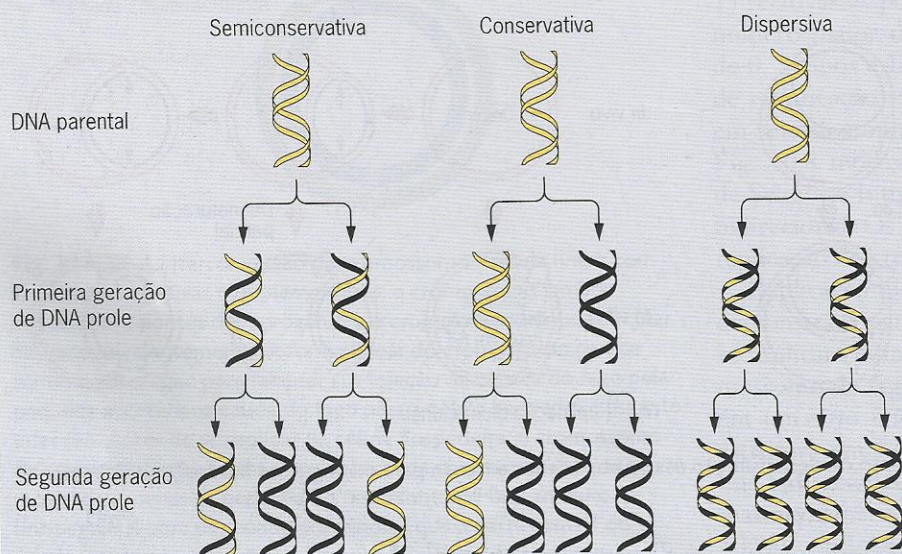


Fig. 1 ■ Três modos possíveis de replicação do DNA: (1) semiconservativa, na qual cada filamento da dupla hélice parental é conservado e dirige a síntese de um novo filamento prole complementar; (2) conservativa, na qual a dupla hélice parental é conservada e dirige a síntese de uma nova dupla hélice prole; e (3) dispersiva, na qual os segmentos de cada filamento parental são conservados e dirigem a síntese de novos segmentos complementares que são subsequente unidos para produzir uma prole de novos filamentos.

¹Watson, J. D., e F. H. C. Crick. 1953. Genetical implications of the structure of deoxyribonucleic acid. *Nature* 171:964-969.

²Meselson, M., e F. Stahl. 1958. The replication of DNA in *Escherichia coli*. *Proc. Natl. Acad. Sci., U.S.A.* 44:671-682.

³Cairns, J. 1962. A proof that the replication of DNA involves separation of the strands. *Nature* 194:1274.



UM MARCO NA GENÉTICA (continuação)

DNA “pesado”), lavaram-nas para remover o meio contendo ^{15}N e as transferiram para um meio contendo ^{14}N . Depois que as células se multiplicaram na presença de ^{14}N por vários períodos de tempo, os DNA foram extraídos e analisados em gradientes de densidade de equilíbrio de CsCl . Os resultados de seu experimento (Fig. 2) são consistentes apenas com a replicação semiconservativa, excluindo os modelos conservativo e dispersivo de síntese de DNA. Todo o DNA isolado das células após uma geração de crescimento em meio contendo ^{14}N tinha uma densidade que era a metade entre as densidades do DNA “pesado” e do DNA “leve”. Essa densidade intermediária geralmente é chamada densidade “híbrida”. Após duas gerações de crescimento em meio contendo ^{14}N , metade do DNA era de densidade híbrida e metade era leve. Tais resultados são exatamente os previstos pelo modelo semiconservativo de replicação

de Watson e Crick (Fig. 2). Uma geração de replicação semiconservativa de uma dupla hélice parental contendo ^{15}N no meio contendo apenas ^{14}N produziria duas hélices duplas, ambas tendo ^{15}N em um filamento (o filamento “antigo”) e ^{14}N no outro filamento (o filamento “novo”). Tais moléculas seriam de densidade híbrida.

A replicação conservativa não produziria moléculas de DNA com densidade híbrida. Após uma geração de replicação conservativa de DNA pesado em meio leve, metade do DNA ainda seria pesado e metade seria leve. Se a replicação fosse dispersiva, Meselson e Stahl teriam observado uma mudança do DNA de pesado para leve em cada geração (isto é, “metade pesado” ou híbrido após uma geração, “um quarto pesado” após duas gerações e assim em diante). Estas possibilidades são claramente inconsistentes com os resultados do experimento de Meselson e Stahl.

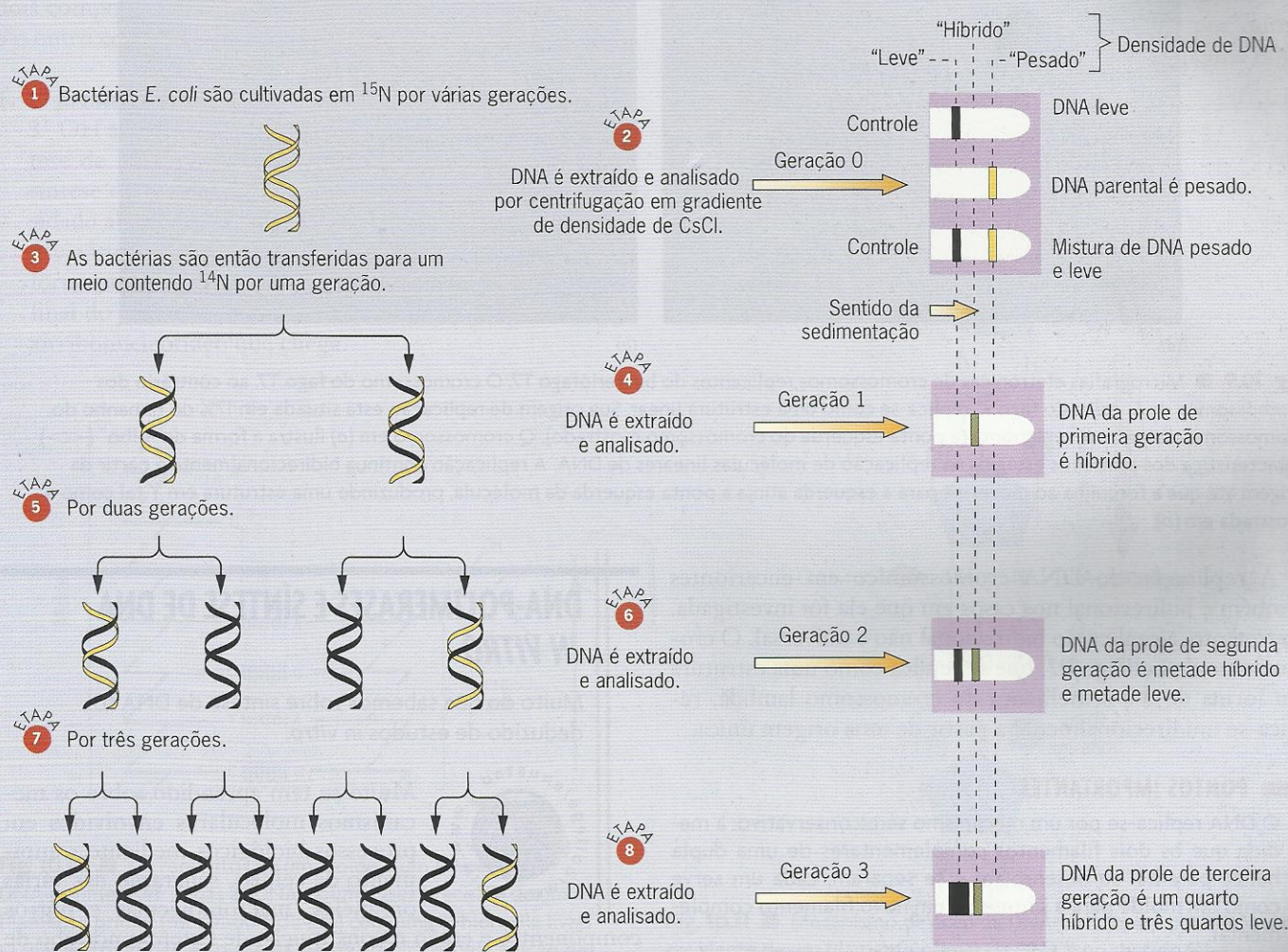


Fig. 2 ■ Experimento de Meselson e Stahl demonstrando que o DNA se replica por um mecanismo semiconservativo em *E. coli*. O diagrama mostra que os resultados de seu experimento são os esperados se o cromossomo de *E. coli* replica-se semiconservativamente. Resultados diferentes teriam sido obtidos se a replicação do DNA em *E. coli* fosse ou conservativa ou dispersiva (veja Fig. 1).



UM MARCO NA GENÉTICA (continuação)

QUESTÕES PARA DISCUSSÕES

1. A replicação do DNA deve ser altamente precisa para facilitar a transferência de informação genética dos genitores para a prole. Entretanto, ela não pode ser muito precisa porque é necessária nova variação genética para que ocorra evolução, permitindo que os organismos se adaptem a mudanças no meio. O que determina o equilíbrio apropriado entre a transmissão precisa do material genético de genitores para a prole e a geração de nova variação, isto é, a ocorrência de novas mutações?
2. Uma cópia do genoma humano contém cerca de 3 bilhões de pares de nucleotídeos. Apenas um pouco mais de 1% desse DNA codifica proteínas, enquanto as partes não codificantes de genes constituem cerca de 24% do genoma. Assim, quase 75% do genoma humano é de DNA intergênico. Considerando que se gasta muita energia para replicar esse DNA não codificante cada vez que uma célula se divide, por que esse DNA é mantido no genoma? Ele desempenha funções importantes? Que funções pode desempenhar um DNA não codificante? Algumas de tais sequências intergênicas são altamente conservadas em espécies correlatas. Essa conservação fornece alguma informação sobre sua importância?



(a)



(b)

Fig. 10.9 ■ Micrografias eletrônicas de cromossomos replicantes do bacteriófago T7. O cromossomo do fago T7, ao contrário dos cromossomos de *E. coli* e do fago λ , replica-se como uma estrutura linear. Sua origem de replicação está situada em 17% do tamanho do cromossomo distante de uma ponta (a ponta esquerda do cromossomo mostrado). O cromossomo em (a) ilustra a forma de “olho” (—○—) característica dos primeiros estágios na replicação de moléculas lineares de DNA. A replicação continua bidirecionalmente a partir da origem até que a forquilha ao mover-se para a esquerda atinja a ponta esquerda da molécula, produzindo uma estrutura em Y tal como a mostrada em (b).

A replicação do DNA cromossômico em eucariontes também é bidirecional nos casos em que ela foi investigada. Entretanto, a replicação bidirecional não é universal. O cromossomo do colifago P2, que se replica como uma estrutura em forma de θ à semelhança do cromossomo lambda, replica-se unidirecionalmente a partir de uma origem única.

PONTOS IMPORTANTES

- O DNA replica-se por um mecanismo semiconservativo: à medida que os dois filamentos complementares de uma dupla hélice parental se desenrolam e se separam, cada um serve como um molde para a síntese de um novo filamento complementar.
- Os potenciais de pontes de hidrogênio das bases nos filamentos moldes especificam as sequências de bases complementares nos filamentos de DNA nascentes.
- A replicação é iniciada em origens únicas e em geral continua bidirecionalmente a partir de cada origem.

DNA-POLIMERASES E SÍNTESE DE DNA IN VITRO

Muito do que sabemos sobre síntese de DNA foi deduzido de estudos *in vitro*.



Muito se tem aprendido sobre os mecanismos moleculares envolvidos em processos biológicos mediante rompimento de células, separação das várias organelas, macromoléculas e outros componentes e então reconstituição de sistemas no tubo de ensaio, chamados sistemas *in vitro*, que são capazes de efetuar determinados eventos metabólicos. Tais sistemas *in vitro* podem ser dissecados bioquimicamente com muito mais facilidade do que sistemas *in vivo*. Claramente, a informação obtida de estudos de sistemas *in vitro* tem sido inestimável.

vel. Entretanto, nunca devemos supor que um fenômeno demonstrado *in vitro* ocorre *in vivo*. Tal extrapolação só deve ser feita quando evidências independentes de estudos *in vivo* validam os estudos *in vitro*.

Descoberta da DNA-polimerase I em *Escherichia coli*

A síntese *in vitro* de DNA foi primeiro feita por Arthur Kornberg e seus colaboradores em 1957. Kornberg, que recebeu um Prêmio Nobel em 1959 por seu trabalho, isolou uma enzima de *E. coli* que catalisa a adição covalente de nucleotídeos a cadeias de DNA preexistentes. Inicialmente chamada DNA-polimerase ou “enzima de Kornberg”, hoje é conhecida como **DNA-polimerase I**. A enzima requer os 5'-trifosfatos de cada um dos quatro desoxirribonucleosídeos – desoxiadenosina-trifosfato (dATP), desoxitimidina-trifosfato (dTTP), desoxiguanosina-trifosfato (dGTP) e desoxicitidina-trifosfato (dCTP) – e só é ativa na presença de íons Mg^{2+} e DNA preexistente. Esse DNA deve fornecer dois componentes essenciais, um servindo como um *primer* e o outro como um molde (Fig. 10.10).

1. O **primer de DNA** fornece um término com uma ponta 3'-OH à qual são adicionados nucleotídeos durante a síntese de DNA. A DNA-polimerase I não pode iniciar a síntese de cadeias de DNA *de novo*. Ela tem uma necessidade absoluta de uma 3'-hidroxila livre em uma cadeia preexistente de DNA. A DNA-polimerase I catalisa a formação de uma ponte fosfodiéster entre a 3'-OH no final do *primer* da cadeia de DNA e o 5'-fosfato do desoxirribonucleotídeo que chega.

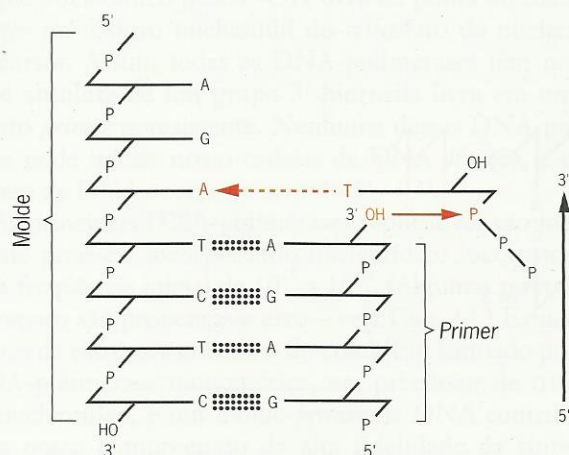


Fig. 10.10 ■ Os requisitos de um molde e um *primer* das DNA-polimerases. Todas as DNA-polimerases precisam de um filamento *primer* (mostrado à direita) com uma 3'-hidroxila livre. O filamento *primer* é covalentemente ampliado pela adição de nucleotídeos (tais como dTMP, derivado do precursor dTTP que chega). Além disso, as DNA-polimerases precisam de um filamento molde (mostrado à esquerda), que determina a sequência de bases do filamento que está sendo sintetizado. O novo filamento será complementar ao filamento molde.

2. O **molde de DNA** fornece a sequência de nucleotídeos que especifica a sequência complementar da cadeia nascente de DNA. A DNA-polimerase I requer um molde de DNA cuja sequência de bases dite, por seu potencial de pareamento de bases, a síntese de uma sequência de bases complementares no filamento a ser sintetizado.

A reação catalisada pela DNA-polimerase I é um ataque nucleofílico pela 3'-OH à ponta do filamento *primer* no nucleotídeo ou átomo interior de fósforo do nucleosídeo trifosfato precursor com a eliminação de pirofosfato. Este mecanismo de reação explica a necessidade absoluta de DNA-polimerase I de um grupo 3'-OH livre no filamento *primer* do DNA que está sendo covalentemente ampliado e determina que o *sentido da síntese seja sempre de 5' → 3'* (Fig. 10.11).

A DNA-polimerase I é um polipeptídeo único com peso molecular de 103.000 codificado por um gene chamado *polA*. Entretanto, pesquisas subseqüentes mostraram que a DNA-polimerase I não é a verdadeira “DNA-replicase” em *E. coli*. Ela não catalisa a replicação semiconservativa do cromossomo de *E. coli*. Esta função é feita por outra enzima. Entretanto, a DNA-polimerase I desempenha fun-

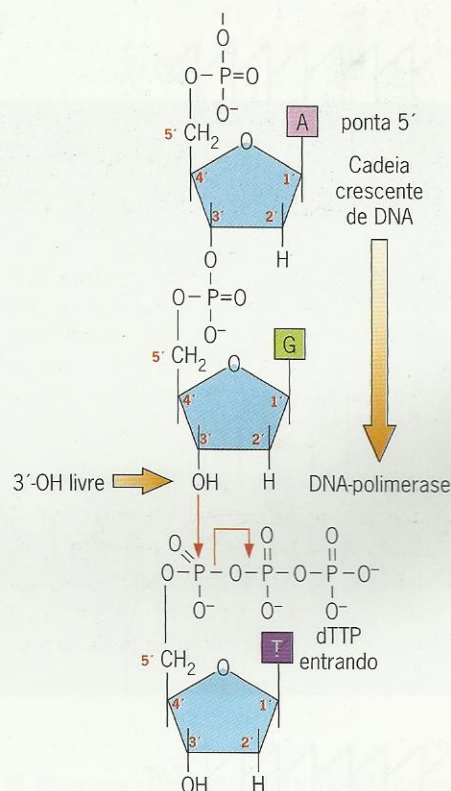


Fig. 10.11 ■ Mecanismo de ação da DNA-polimerase I: a ampliação covalente de um filamento *primer* de DNA é no sentido 5' → 3'. A cadeia existente termina na ponta 3' com o nucleotídeo desoxiguanilato (desoxiguanosina-5'-fosfato). O diagrama mostra a adição catalisada pela DNA-polimerase do monofosfato de desoxitimidina (a partir do precursor trifosfato de desoxitimidina, dTTP) na ponta 3' da cadeia com a liberação de pirofosfato (P_2O_7).

ções importantes na *E. coli*, incluindo um papel principal na replicação do cromossomo e um papel central na reparação do DNA danificado. Para entender como a DNA-polimerase I desempenha tais funções, precisamos conhecer melhor esta enzima.

Além da atividade de polimerase ilustrada na Fig. 10.11, a DNA-polimerase I tem duas outras atividades enzimáticas, ambas atividades de exonuclease. Uma **nuclease** é uma enzima que degrada ácidos nucleicos. Uma **exonuclease** degrada ácidos nucleicos começando em uma ou ambas as pontas, enquanto uma **endonuclease** corta ácidos nucleicos em sítios internos. A DNA-polimerase I contém tanto **atividade 5' → 3' de exonuclease**, que corta filamentos de DNA começando pelas pontas 5', quanto **atividade de 3' → 5' de exonuclease**, que corta mononucleotídeos das pontas 3' dos filamentos de DNA. A atividade de exonuclease 5' → 3' da DNA-polimerase I em geral remove pequenos oligômeros contendo até 10 nucleotídeos. Assim, a DNA-polimerase I exerce três atividades enzimáticas diferentes: (1) uma atividade de polimerase 5' → 3', (2) uma atividade de exonuclease 5' → 3' e (3) uma atividade de exonuclease 3' → 5'. Estas três atividades são ilustradas na Fig. 10.12.

A primeira evidência de que a DNA-polimerase I não era a verdadeira DNA-replicase foi publicada em 1969 por Paula DeLucia e John Cairns. Eles relataram que a replicação do DNA ocorreu em uma linhagem de *E. coli* sem atividade de polimerase 5' → 3' (mas não a atividade de exonuclease 5' → 3') desta enzima devido a uma mutação no gene *polA*. Seus resultados demonstraram que a replicação do DNA em *E. coli* não requer atividade da DNA-polimerase I, pelo menos não a atividade da polimerase 5' → 3' da enzima. Entretanto, DeLucia e Cairns também descobriram que este mutante *polA1* era extremamente sensível a luz ultravioleta (UV). Todas as três atividades enzimáticas da DNA-polimerase I têm papéis importantes na célula. A principal função da DNA-polimerase I em *E. coli* é reparar defeitos no DNA, tais como os induzidos por UV (Cap. 13). Entretanto, veremos mais adiante neste capítulo que a atividade 5' → 3' de exonuclease da DNA-polimerase I está também envolvida em um estágio de replicação cromossômica.

Múltiplas DNA-polimerases

Se a DNA-polimerase I não catalisa a replicação semiconservativa do cromossomo de *E. coli*, outra polimerase deve exer-

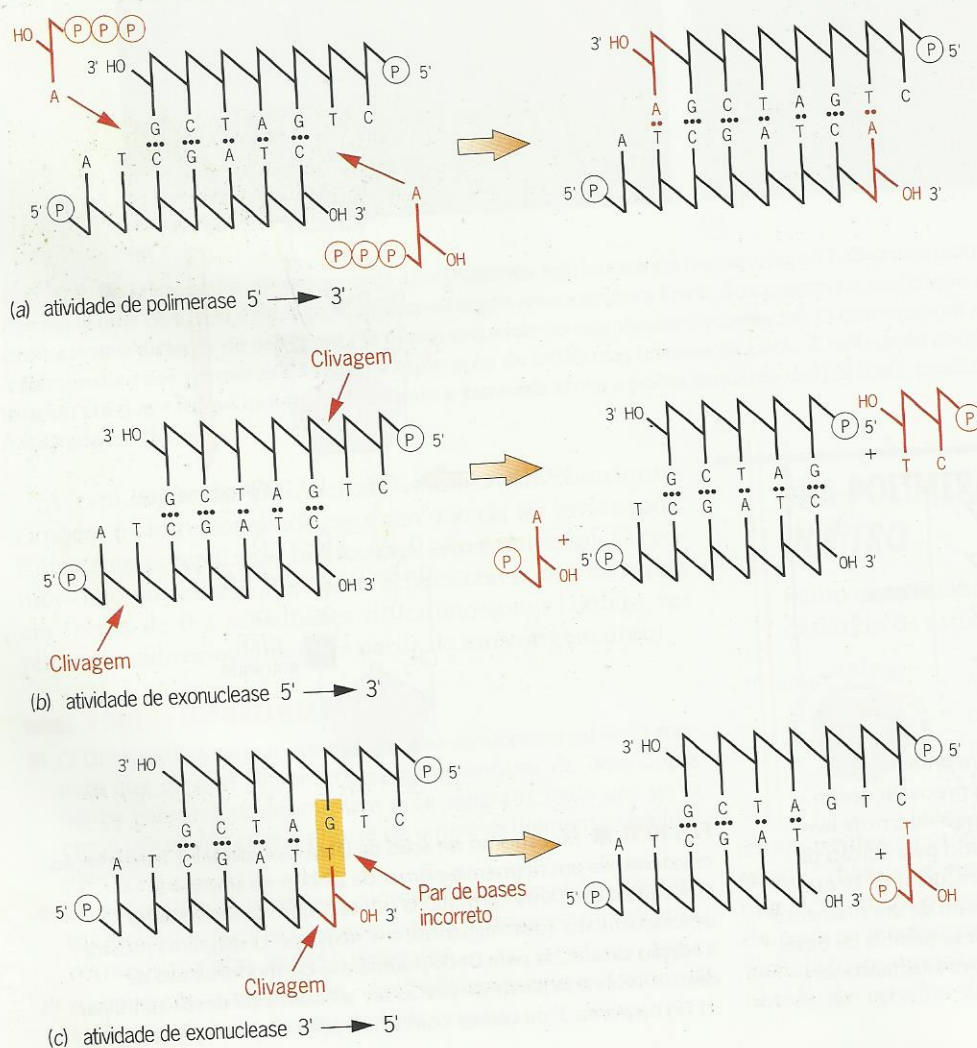


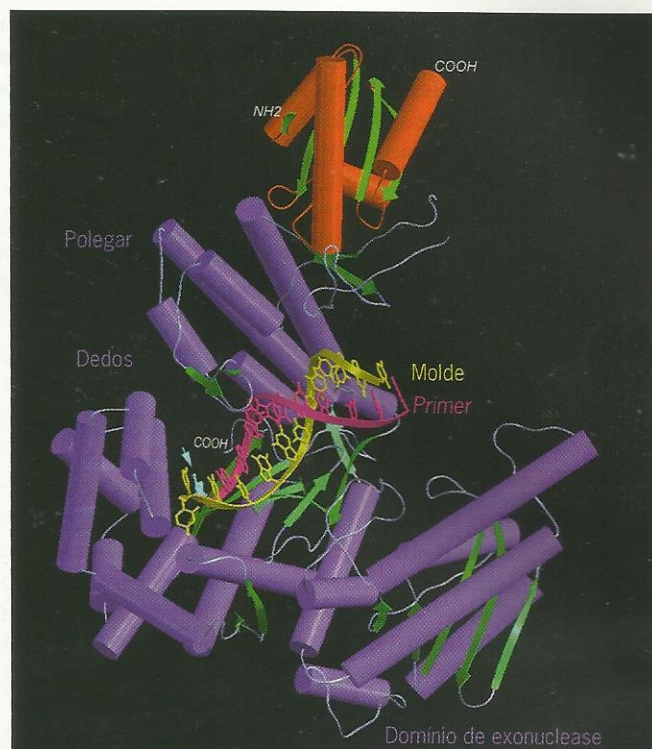
Fig. 10.12 ■ As três atividades da DNA-polimerase I de *E. coli*. Como discutido no texto, todas as três atividades têm papéis importantes na replicação do cromossomo de *E. coli*.

cer esta função. De fato, existem pelo menos quatro outras DNA-polimerases, **DNA-polimerase II**, **DNA-polimerase III**, **DNA-polimerase IV** e **DNA-polimerase V**, em *E. coli*. Como a DNA-polimerase I, a DNA-polimerase II é uma enzima de reparo do DNA, mas ela representa uma pequena proporção da atividade de polimerase em uma *E. coli*. A DNA-polimerase II é um polipeptídeo único com atividades de polimerase $5' \rightarrow 3'$ e exonuclease $3' \rightarrow 5'$. Entretanto, ela não exerce atividade de exonuclease $5' \rightarrow 3'$. Em contraste com as DNA-polimerases I e II, as DNA-polimerases III são um complexo enzimático composto de muitas subunidades diferentes. Como a DNA polimerase II, a DNA polimerase III exerce atividades de polimerase $5' \rightarrow 3'$ e exonuclease $3' \rightarrow 5'$. Entretanto, ela tem uma exonuclease $5' \rightarrow 3'$ que é ativa apenas no DNA unifilar. As DNA-polimerases IV e V mais recentemente caracterizadas, juntamente com a polimerase II, desempenham papéis importantes na replicação de DNA danificado, com a polimerase envolvida dependendo do tipo de dano. Suas funções são mais bem discutidas no Cap. 13.

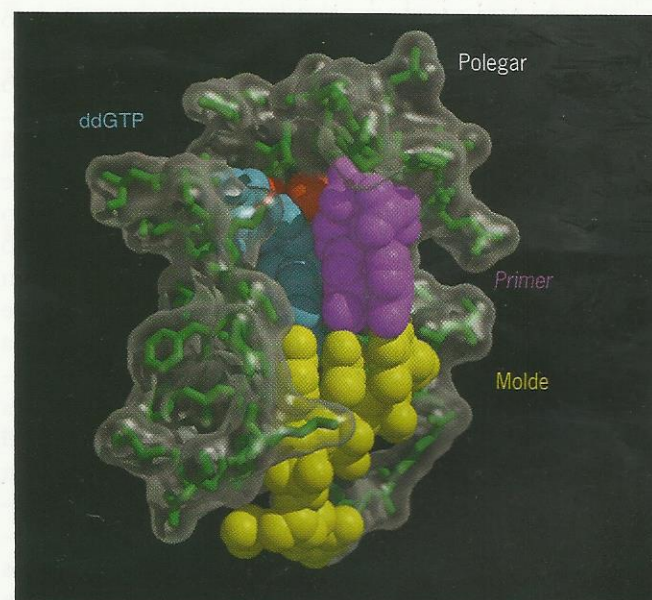
Organismos eucarióticos são ainda mais complexos, com sete DNA-polimerases diferentes tendo sido identificadas até agora. As sete DNA-polimerases caracterizadas em mamíferos foram denominadas α , β , γ , δ , ϵ , ζ e η . Duas ou mais DNA-polimerases (α , δ , e/ou ϵ) funcionam juntas para efetuar a replicação semiconservativa do DNA nuclear. A DNA-polimerase γ é responsável pela replicação do DNA em mitocôndrias e as DNA-polimerases β , ζ e η são enzimas nucleares de reparo do DNA. Algumas das DNA-polimerases de eucariontes não têm atividade de exonuclease $3' \rightarrow 5'$ que está presente em DNA-polimerases procarióticas.

Todas as DNA-polimerases estudadas até agora, procarióticas e eucarióticas, catalisam a mesma reação básica: um ataque nucleofílico pela $3'$ -OH livre na ponta do filamento *primer* no fósforo nucleotídil do trifosfato de nucleosídeo precursor. Assim, todas as DNA-polimerases têm necessidade absoluta de um grupo $3'$ -hidroxila livre em um filamento *primer* preexistente. Nenhuma dessas DNA-polimerases pode iniciar novas cadeias de DNA *de novo*, e toda a síntese de DNA ocorre no sentido $5' \rightarrow 3'$.

As principais DNA-polimerases replicativas são incrivelmente precisas, incorporando nucleotídeos incorretos com uma frequência inicial de 10^{-5} a 10^{-6} . (Algumas polimerases de reparo são propensas a erro – veja Cap. 13.) Estudos recentes da estrutura cristalina do complexo formado por uma DNA-polimerase monomérica, um precursor de trifosfato de nucleosídeo, e um molde-*primer* de DNA contribuíram para nossa compreensão da alta fidelidade da síntese de DNA. Nesses estudos, publicados em 1998, Sylvie Doublié e colaboradores determinaram a estrutura da polimerase do fago T7, que é similar à DNA-polimerase I de *E. coli*, com uma resolução de 0,22 nm. Os resultados mostram que a polimerase tem a forma de uma pequena mão, com o trifosfato de nucleosídeo que chega, o molde, e o terminal do *primer* todos fortemente fixados entre o polegar, os dedos e a palma (Fig. 10.13). A enzima posiciona o trifosfato de nucleosídeo que chega em justaposição com o término do filamento *pri-*



(a)



(b)

Fig. 10.13 ■ Diagrama esquemático (a) e modelo compacto (b) da estrutura do complexo entre a DNA-polimerase do fago T7, o *primer* do DNA molde, e uma molécula precursora do trifosfato de nucleosídeo (ddGTP). O filamento molde, filamento *primer*, e trifosfato de nucleosídeo são mostrados em amarelo, magenta e ciano, respectivamente. Os componentes protéicos são mostrados em púrpura, verde, laranja e cinza. Note a justaposição íntima entre o trifosfato de nucleosídeo, o terminal do *primer* e o filamento molde em (b).

mer em uma posição para formar pontes de hidrogênio com a primeira base não pareada no filamento molde. Assim, a estrutura deste complexo polimerase fornece uma explicação simples para a seleção dirigida por molde dos nucleotídeos que chegam durante a síntese de DNA.

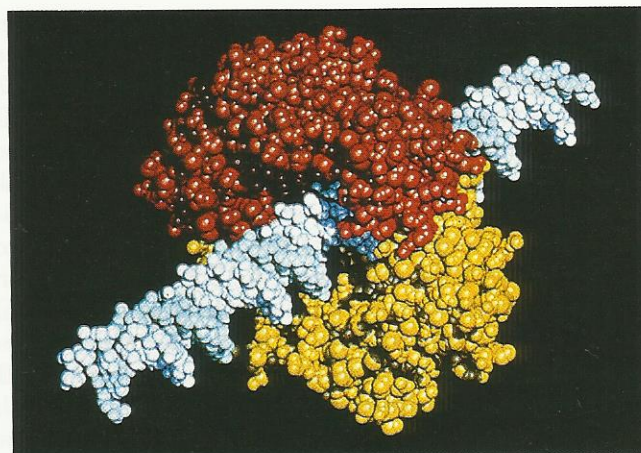
DNA-polimerase III: A Replicase em *Escherichia coli*

Evidência indicativa de que a DNA-polimerase III é a verdadeira DNA-replicase responsável pela replicação semiconservativa do DNA em *E. coli* foi primeiro obtida isolando-se e caracterizando-se uma linhagem mutante com uma mutação em um gene chamado *polC*, hoje com a nova designação de *dnaE*. Esta linhagem mutante produziu uma DNA-polimerase III ativa quando cultivada a 25°C, mas uma polimerase III totalmente inativa quando cultivada a 43°C. Quando bactérias mutantes *dnaE* cultivadas a 25°C foram mudadas para 43°C, a replicação do DNA cessou, indicando que o produto do gene *dnaE* é necessário para síntese de DNA. Foi subsequentemente demonstrado que o gene *dnaE* codifica a subunidade α catalítica, a subunidade com atividade de polimerase 5' \rightarrow 3', da DNA-polimerase III.

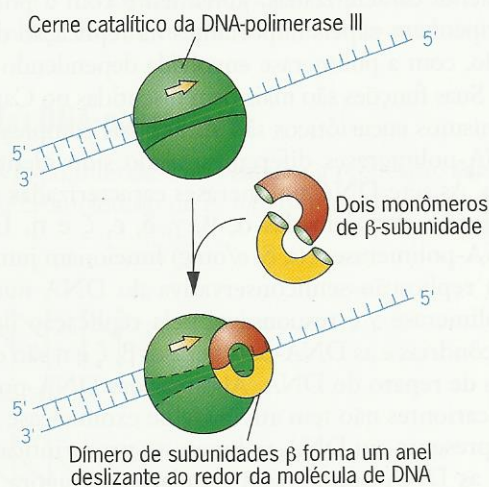
A DNA-polimerase III é uma enzima multimérica (uma enzima com muitas subunidades) com massa molecular de cerca de 900.000 dáltons em sua forma completa, ou **holoenzima**. O cerne mínimo que tem atividade catalítica *in vitro* contém três subunidades: α (o produto do gene *dnaE*), ϵ (o produto do gene *dnaQ*) e θ (o produto de *holE*). A adição da subunidade τ (o produto do *dnaX*) resulta em dimerização do cerne catalítico e atividade aumentada. O cerne catalítico sintetiza filamentos de DNA bem curtos devido a sua tendência a sair do molde de DNA. Para produzir as moléculas longas de DNA presentes em cromossomos, essa freqüente dissociação da polimerase do molde deve ser eliminada. A subunidade β (o produto do gene *dnaN*) da DNA-polimerase III forma um grampo dimérico que impede a polimerase de sair do molde de DNA (Fig. 10.14). O β -dímero forma um anel que rodeia a molécula de DNA replicante e permite que a DNA-polimerase III deslize ao longo do DNA permanecendo amarrada a ele. A holoenzima DNA-polimerase III, que é responsável pela síntese de ambos os filamentos nascentes do DNA em uma forquilha de replicação, contém pelo menos 20 polipeptídeos. A complexidade estrutural da holoenzima DNA-polimerase III é ilustrada na Fig. 10.15; o diagrama mostra 16 dos polipeptídeos mais bem caracterizados codificados por sete genes diferentes. Consideraremos as funções de algumas das subunidades em seções seguintes deste capítulo.

Atividades de Revisão das DNA-polimerases

Como já discutimos, a fidelidade da duplicação do DNA é incrível, com apenas um erro presente em cada bilhão de pares de bases logo após a síntese. Essa alta fidelidade é necessária para manter a carga de mutação em nível tolerável, especialmente em genomas grandes tais como os de mamíferos, que



(a)



(b)

Fig. 10.14 ■ Modelo compacto (a) e diagrama (b) mostrando como duas subunidades β (vermelha e amarela) da DNA-polimerase III prendem a enzima à molécula de DNA (azul).

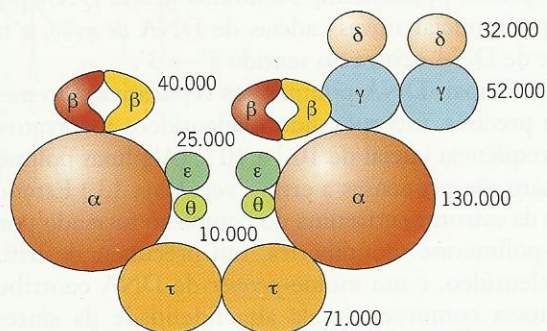


Fig. 10.15 ■ Estrutura da holoenzima DNA-polimerase III de *E. coli*. Os números dão a massa em dáltons.

contêm 3×10^9 pares de nucleotídeos. Sem a alta fidelidade da replicação do DNA, os gêmeos monozygóticos discutidos no começo deste capítulo teriam fenótipos menos similares. De fato, com base nas estruturas dinâmicas de quatro nucleotídeos no DNA, a fidelidade observada na replicação do DNA é

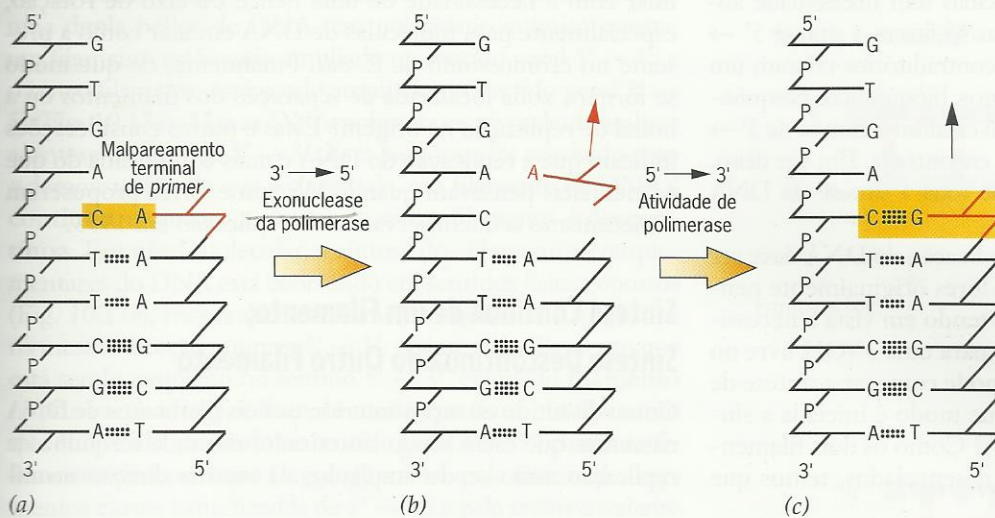


Fig. 10.16 ■ Revisão pela atividade de exonuclease 3' → 5' das DNA-polimerases durante replicação do DNA. Se a DNA-polimerase está presente com um molde e um primer contendo um pareamento errado terminal em um primer 3' (a), ela não irá catalisar a extensão covalente (polimerização). Em vez disso, a atividade de exonuclease 3' → 5', uma parte integral de muitas DNA-polimerases, eliminará os nucleotídeos malpareados terminais (b). Então, com o terminal de um primer com pares de bases corretas, a DNA-polimerase irá catalisar uma extensão covalente 5' → 3' do filamento primer (c).

muito maior que a esperada. As mudanças termodinâmicas nos nucleotídeos que permitem a formação de pontes de hidrogênio entre pares de bases à exceção de A-T e G-C prevêm taxas de erro de 10^{-5} a 10^{-4} , ou um erro por 10.000 a 100.000 nucleotídeos incorporados. A taxa de erro prevista de 10.000 vezes a taxa observada levanta a dúvida sobre como é obtida essa alta fidelidade de replicação do DNA.

Os organismos vivos resolveram o problema potencial de fidelidade insuficiente durante replicação do DNA desenvolvendo um mecanismo de **revisão** da cadeia nascente de DNA à medida que ele está sendo produzido. O processo de revisão envolve escanear as pontas das cadeias nascentes de DNA quanto a erros e corrigi-los. Este processo é feito pelas atividades de 3' → 5' de exonuclease das DNA-polimerases. Quando um molde-primer de DNA tem um malpareamento terminal (uma base ou sequência de bases pareada incorretamente ou não pareada na ponta 3' do primer), a atividade 3' → 5' de exonuclease da DNA-polimerase remove a base ou bases não pareadas (Fig. 10.16). Quando é produzido um terminal com pares de bases apropriadas, a atividade 5' → 3' de polimerase da enzima começa nova síntese adicionando nucleotídeos à ponta 3' do filamento primer. Em enzimas monoméricas como a DNA-polimerase I de *E. coli*, essa atividade está incluída. Em enzimas multiméricas, a atividade 3' → 5' de revisão da exonuclease em geral está presente em uma subunidade separada. No caso da DNA-polimerase III de *E. coli*, esta função de revisão é feita pela subunidade ϵ . A DNA-polimerase IV de *E. coli* não exerce atividade de exonuclease. Em eucariontes, as DNA-polimerases γ , δ e ϵ exercem atividades de exonuclease de revisão 3' → 5', mas as polimerases α e β não têm esta atividade. Considerando a importância da revisão, podemos especular que proteínas acessórias devem efetuar a função de revisão para as DNA-polimerases α e β .

Sem a revisão durante replicação do DNA, Merry e Sherry, as gêmeas discutidas no começo deste capítulo, teriam aspectos menos similares. Sem a revisão, as mudanças teriam se acumulado em seus genes durante os bilhões de divisões celulares que sucederam no decorrer de seu crescimento desde pequenos embriões até adultos. De fato, a

identidade dos genótipos de gêmeos idênticos depende tanto de revisão do DNA durante a replicação quanto da atividade em um exército de enzimas de reparo do DNA (Cap. 13). Tais enzimas percorrem continuamente o DNA quanto a vários tipos de danos e fazem reparos antes que as alterações causem mudanças genéticas herdadas.

PONTOS IMPORTANTES

- A síntese de DNA é catalisada por enzimas chamadas DNA-polimerases.
- Todas as DNA-polimerases precisam de um filamento *primer*, que é ampliado, e um filamento molde, que é copiado.
- Todas as DNA-polimerases têm necessidade absoluta de uma 3'-OH livre no filamento *primer*, e toda a síntese de DNA ocorre no sentido 5' para 3'.
- As atividades de exonuclease 3' → 5' das DNA-polimerases revisam filamentos nascentes à medida que eles são sintetizados, removendo quaisquer nucleotídeos malpareados nas pontas 3' dos filamentos *primer*.

O COMPLEXO APARELHO DE REPLICAÇÃO

A replicação do DNA é um processo complexo, que requer a ação conjunta de um grande número de proteínas.



Os resultados de estudos de replicação do DNA por auto-radiografia e microscopia eletrônica indicam que os dois filamentos prole que estão sendo sintetizados em cada forquilha de replicação estão sendo estendidos no mesmo sentido geral em nível macromolecular. Como os filamentos complementares de uma dupla hélice têm polaridade oposta, a síntese está ocorrendo na ponta 5' de um filamento (extensão 3' → 5') e na ponta 3' do outro filamento (extensão 5' → 3'). Entretanto, como já discutimos anterior-

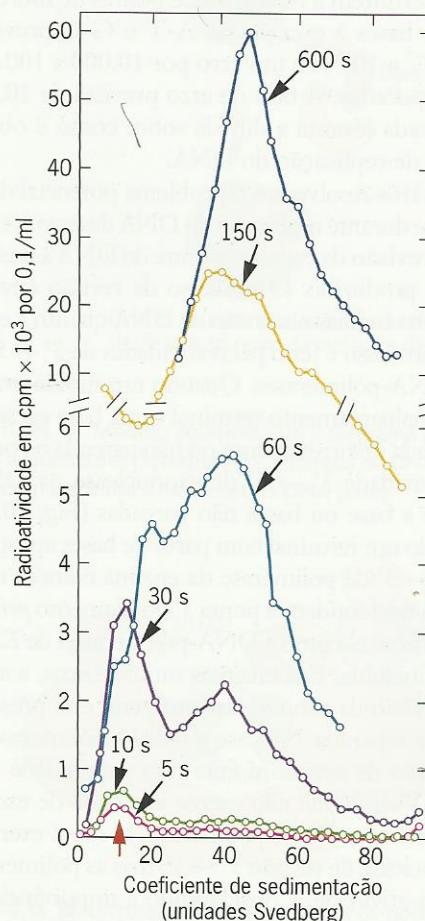
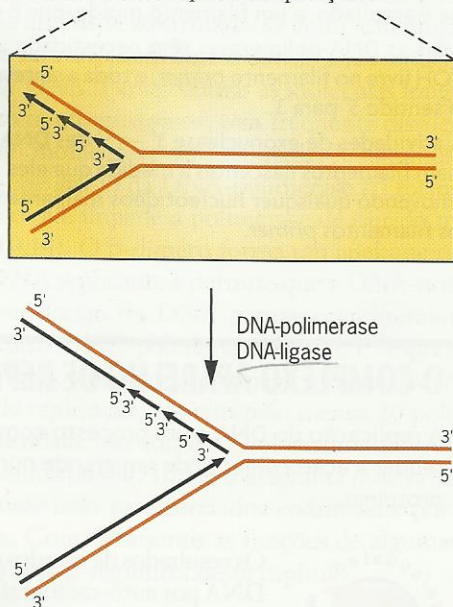
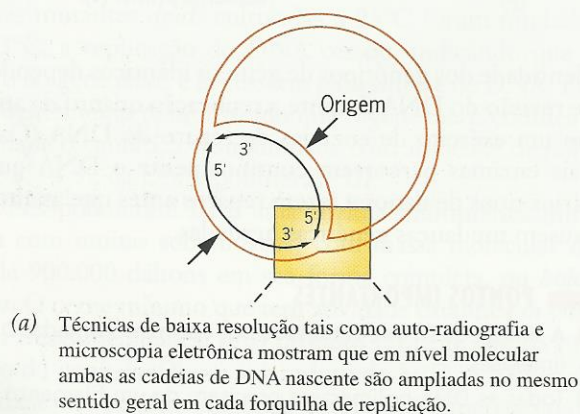
mente, todas as polimerases conhecidas têm necessidade absoluta de uma 3'-hidroxila livre. Elas só fazem a síntese $5' \rightarrow 3'$. Estes resultados aparentemente contraditórios criaram um paradoxo interessante. Por muitos anos, bioquímicos pesquisaram novas polimerases que pudessem catalisar a síntese de $3' \rightarrow 5'$. Nenhuma dessas polimerases foi encontrada. Em vez disso, evidência experimental mostrou que toda a síntese de DNA ocorre no sentido $5' \rightarrow 3'$.

Claramente, o mecanismo de replicação do DNA deve ser mais complexo do que os pesquisadores originalmente pensavam (veja Fig. 10.1). Além disso, tendo em vista a necessidade absoluta da DNA-polimerase para uma 3'-OH livre no filamento *primer*, esta enzima não pode começar a síntese de um novo filamento de novo. De que modo é iniciada a síntese de um novo filamento de DNA? Como os dois filamentos parentais de DNA devem ser desenrolados, temos que

lidar com a necessidade de uma hélice ou eixo de rotação, especialmente para moléculas de DNA circular como a presente no cromossomo de *E. coli*. Finalmente, de que modo se forma a zona localizada de separação dos filamentos ou a bolha de replicação na origem? Estas e outras considerações indicam que a replicação do DNA é mais complicada do que os cientistas pensavam quando Watson e Crick propuseram o mecanismo semiconservativo de replicação em 1953.

Síntese Contínua de um Filamento; Síntese Descontínua do Outro Filamento

Como discutido na seção anterior, os dois filamentos de DNA nascentes que estão sendo sintetizados em cada forquilha de replicação estão sendo ampliados na mesma direção no ní-



(c) Análise de velocidade de gradiente de sacarose de DNA de *E. coli* com marcação de pulso de ^3H -timidina, extraído e desnaturado durante centrifugação.

Fig. 10.17 ■ Evidência de síntese descontínua do filamento *lagging*. (a) Embora ambos os filamentos do DNA nascente sintetizado em uma forquilha de replicação pareçam ser ampliados no mesmo sentido, (b) em nível molecular, eles de fato são sintetizados em sentidos opostos. (c) Resultados de experimentos de marcação de pulso de Reiji e Tuneko Okazaki e colaboradores mostrando que o DNA nascente em *E. coli* existe em pequenos fragmentos de 1.000 a 2.000 nucleotídeos de tamanho. A seta vermelha mostra a posição dos "fragmentos de Okazaki" no gradiente.

vel macromolecular. Como os filamentos complementares de uma dupla hélice de DNA têm polaridade química oposta, um filamento está sendo ampliado no sentido geral $5' \rightarrow 3'$, e o outro filamento está sendo ampliado no sentido geral $3' \rightarrow 5'$ (Fig. 10.17a). Mas as DNA-polimerases só podem catalisar a síntese no sentido $5' \rightarrow 3'$. Este paradoxo foi resolvido com a demonstração de que a síntese de um filamento de DNA é **contínua**, enquanto a síntese do outro filamento é **descontínua**. Em nível molecular, a síntese dos filamentos complementares do DNA está ocorrendo em sentidos físicos opostos (Fig. 10.17b), mas ambos os filamentos novos são ampliados na mesma direção química $5' \rightarrow 3'$. A síntese do filamento que está sendo ampliado no sentido $5' \rightarrow 3'$, chamado **filamento contínuo** (*leading*), é feita de modo **contínuo**. O filamento que está sendo ampliado no sentido geral $3' \rightarrow 5'$, chamado **filamento descontínuo** (*lagging*), cresce pela síntese de fragmentos curtos (sintetizados de $5' \rightarrow 3'$) e pela união covalente subsequente destes fragmentos curtos. Assim, a síntese do filamento *lagging* ocorre por um mecanismo descontínuo.

A primeira evidência deste modo descontínuo de replicação do DNA veio de estudos nos quais os intermediários na síntese de DNA foram marcados radioativamente cultivando-se bactérias *E. coli* e tais bactérias infectadas pelo bacteriófago T4 por curtos períodos de tempo em meio contendo ^3H -timidina (experimentos de pulso-marcação). Os DNA marcados foram então isolados, desnaturados e caracterizados medindo-se sua velocidade de sedimentação através de gradientes de sacarose durante centrifugação em alta velocidade (veja Enfoque Técnico: Técnicas de Centrifugação). Quando *E. coli* foram submetidas a marcação de pulso por 5, 10 ou 30 segundos, por exemplo, grande parte da marcação foi encontrada em pequenos fragmentos de DNA, com 1.000 a 2.000 nucleotídeos de tamanho (Fig. 10.17c). Esses pequenos fragmentos de DNA foram chamados fragmentos de Okazaki em homenagem a Reiji Okazaki e Tüeko Okazaki, os cientistas que os descobriram no final dos anos 1960. Em eucariontes, os fragmentos de Okazaki têm apenas 100 a 200 nucleotídeos de tamanho. Quando são usados períodos mais longos de marcação de pulso, mais marcação é recuperada em moléculas grandes de DNA, supostamente do tamanho dos cromossomos de *E. coli* ou fago T4. Se as células forem marcadas com pulso de ^3H -timidina por um curto período de tempo e então transferidas para meio não radioativo por um longo período de crescimento (experimentos de pulso-caça), a timidina marcada está presente em moléculas de DNA do tamanho do cromossomo. Os resultados desses experimentos de pulso-caça são importantes porque indicam que os fragmentos de Okazaki são intermediários verdadeiros na replicação do DNA, e não algum tipo de subproduto metabólico.

Fechamento Covalente dos Cortes no DNA pela DNA-ligase

Se o filamento *lagging* do DNA é sintetizado descontínuamente como descrito na seção anterior, é necessário um meca-

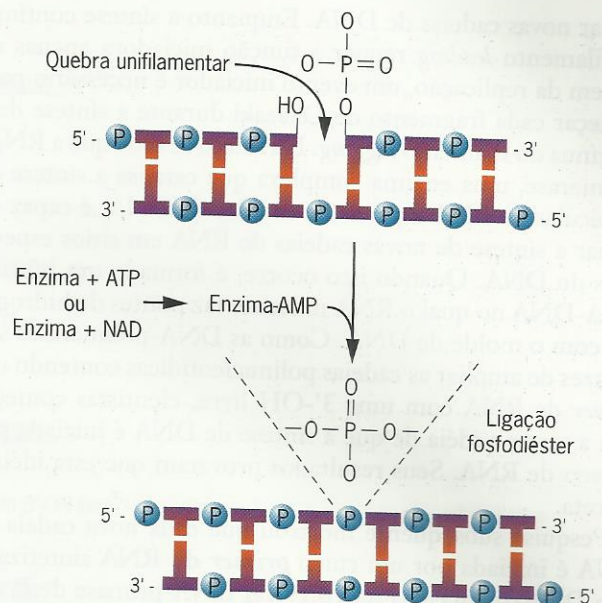


Fig. 10.18 ■ A DNA-ligase catalisa o fechamento covalente de cortes no DNA. A energia necessária para formar ligações éster é fornecida ou por trifosfato de adenosina (ATP) ou por nicotinamida-adenina-dinucleotídeo (NAD), dependendo da espécie.

nismo para unir os fragmentos de Okazaki e produzir os longos filamentos de DNA presentes em cromossomos do produto final. Este mecanismo é feito pela enzima **DNA-ligase**. A DNA-ligase catalisa o fechamento covalente dos cortes (ausência de ligações fosfodiéster; bases presentes) nas moléculas de DNA usando-se energia do dinucleotídeo de nicotinamida e adenina (NAD) ou trifosfato de adenosina (ATP). A DNA-ligase de *E. coli* usa NAD como um cofator, mas algumas DNA-ligasas usam ATP. A reação catalisada pela DNA-ligase é mostrada na Fig. 10.18. Primeiro, o AMP do intermediário ligase-AMP forma uma ligação fosfodiéster com o 5'-fosfato no corte, e então um ataque nucleofílico por 3'-OH no corte do átomo de fósforo proximal do DNA produz uma ligação fosfodiéster entre os nucleotídeos adjacentes no sítio do corte. Apenas a DNA-ligase não tem atividade nas quebras no DNA onde um ou mais nucleotídeos estão faltando, as chamadas falhas (*gaps*). Essas falhas podem ser preenchidas e fechadas apenas pela ação combinada de uma DNA polimerase e uma DNA-ligase. A DNA-ligase desempenha um papel essencial não apenas na replicação do DNA, mas também no reparo e na recombinação do DNA (Cap. 13).

Início das Cadeias de DNA pelos Primers de RNA

Como já foi discutido, todas as DNA-polimerases conhecidas têm necessidade absoluta de uma 3'-OH livre em um filamento *primer* de DNA e um filamento molde de DNA apropriado para atividade. Nenhuma DNA-polimerase conhecida pode iniciar a síntese de um novo filamento de DNA. Assim, deve existir algum mecanismo especial para

iniciar novas cadeias de DNA. Enquanto a síntese contínua do filamento *leading* requer a função iniciadora apenas na origem da replicação, um evento iniciador é necessário para começar cada fragmento de Okazaki durante a síntese descontínua do filamento *lagging*. Há muito se sabe que a RNA-polimerase, uma enzima complexa que catalisa a síntese de moléculas de RNA a partir de moldes de DNA, é capaz de iniciar a síntese de novas cadeias de RNA em sítios específicos do DNA. Quando isto ocorre, é formado um híbrido RNA-DNA no qual o RNA nascente faz pontes de hidrogênio com o molde de DNA. Como as DNA-polimerases são capazes de ampliar as cadeias polinucleotídicas contendo um *primer* de RNA com uma 3'-OH livre, cientistas começaram a testar a idéia de que a síntese de DNA é iniciada por *primers* de RNA. Seus resultados provaram que esta idéia é correta.

Pesquisa subsequente mostrou que cada nova cadeia de DNA é iniciada por um curto **primer de RNA** sintetizado pela **DNA-primase** (Fig. 10.19). A DNA-primase de *E. coli* é o produto do gene *dnaG*. Em procariontes, esses *primers* de RNA têm de 10 a 60 nucleotídeos de tamanho, enquanto em eucariontes eles são menores, com apenas 10 nucleotídeos de tamanho. Os *primers* de RNA fornecem as 3'-OH livres necessárias para ampliação covalente de cadeias polinucleotídicas pelas DNA-polimerases. Em *E. coli*, desoxirribonucleotídeos são adicionados aos *primers* de RNA pela DNA-polimerase III, seja continuamente no filamento *leading* seja descontinuamente pela síntese de fragmentos de Okazaki no filamento *lagging*. A DNA-polimerase III termina um fragmento de Okazaki quando se depara com um *primer* de RNA do fragmento de Okazaki precedente.

Os *primers* de RNA são subsequentemente removidos e substituídos por cadeias de DNA. Esta etapa é feita pela DNA-polimerase I em *E. coli*. Lembre que das cinco DNA-polimerases em *E. coli*, apenas a DNA-polimerase I possui uma exonuclease 5' → 3' que atua na dupla hélice de DNA.

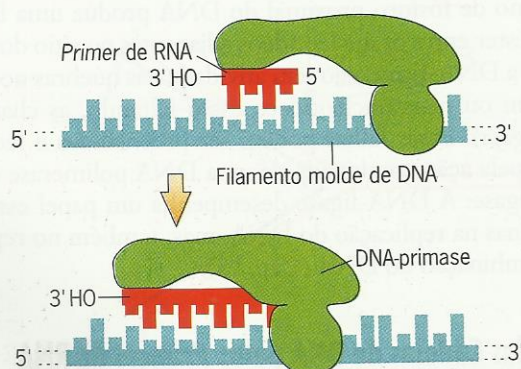


Fig. 10.19 ■ O início dos filamentos de DNA com *primers* de RNA. A enzima DNA-primase catalisa a síntese de curtos (10 a 60 nucleotídeos de tamanho) filamentos de RNA que são complementares aos filamentos moldes. A DNA-polimerase III então usa as 3'-hidroxilas livres dos *primers* de RNA para ampliar as cadeias pela adição de desoxirribonucleotídeos (veja Fig. 10.20).

A atividade de exonuclease 5' → 3' da DNA-polimerase I remove o *primer* de RNA, e, ao mesmo tempo, a atividade de polimerase 5' → 3' da enzima substitui o RNA por uma cadeia de DNA usando o fragmento de Okazaki adjacente com sua 3'-OH livre como um *primer*. Como podíamos esperar com base neste mecanismo de substituição do *primer*, os mutantes *polA* de *E. coli* que não têm atividade de exonuclease 5' → 3' da DNA-polimerase I são defectivos na remoção de *primers* de RNA e na união dos fragmentos de Okazaki. Após a DNA-polimerase I ter substituído o *primer* de RNA por uma cadeia de DNA, a 3'-OH de um fragmento de Okazaki fica em seguida ao grupo 5'-fosfato do fragmento de Okazaki precedente. Este produto é um substrato apropriado para a DNA-ligase, que catalisa a formação de uma ligação fosfodiéster entre os fragmentos de Okazaki adjacentes. As etapas envolvidas na síntese e na substituição dos *primers* de RNA durante a replicação descontínua do filamento *lagging* são ilustradas na Fig. 10.20.

Deselicoização do DNA com Helicases, Proteínas de Ligação ao DNA e Topoisomerases

A replicação semiconservativa requer que os dois filamentos da molécula de DNA parental sejam separados durante a síntese de novos filamentos complementares. Como uma dupla hélice de DNA contém dois filamentos que não podem ser separados sem deselicoizá-los giro a giro, a replicação do DNA requer um mecanismo de desenrolamento. Considerando que cada giro, ou volta, tem cerca de 10 pares de nucleotídeos de tamanho, uma molécula de DNA deve ser girada 360°C para cada 10 pares de bases replicadas. Em *E. coli*, o DNA replica-se a uma velocidade de cerca de 30.000 nucleotídeos por minuto. Assim, uma molécula replicante de DNA deve girar a 3.000 revoluções por minuto para facilitar o desenrolamento dos filamentos parentais de DNA. O processo de desenrolamento (Fig. 10.21a) é catalisado por enzimas chamadas **DNA-helicases**. A principal DNA-helicase replicativa em *E. coli* é o produto do gene *dnaB*. As DNA-helicases desenrolam moléculas de DNA usando energia obtida do ATP.

Uma vez que os filamentos de DNA sejam desenrolados pela DNA-helicase, eles devem ser mantidos na forma unifilar amplificada para replicação. Eles são mantidos neste estado por uma **proteína de ligação a DNA unifilar** (proteína SSB) (Fig. 10.21b). A ligação de proteínas SSB ao DNA unifilar é cooperativa; isto é, a ligação do primeiro monômero SSB estimula a ligação de monômeros adicionais em sítios contíguos na cadeia de DNA. Devido à cooperatividade da proteína de ligação SSB, toda uma região unifilar de DNA é rapidamente revestida por proteína SSB. Sem o revestimento da proteína SSB, os filamentos complementares podem renaturar-se ou formar estruturas de grampos intrafilamentares por meio de pontes de hidrogênio entre pequenos segmentos de seqüências de nucleotídeos complementares ou parcialmente complementares.

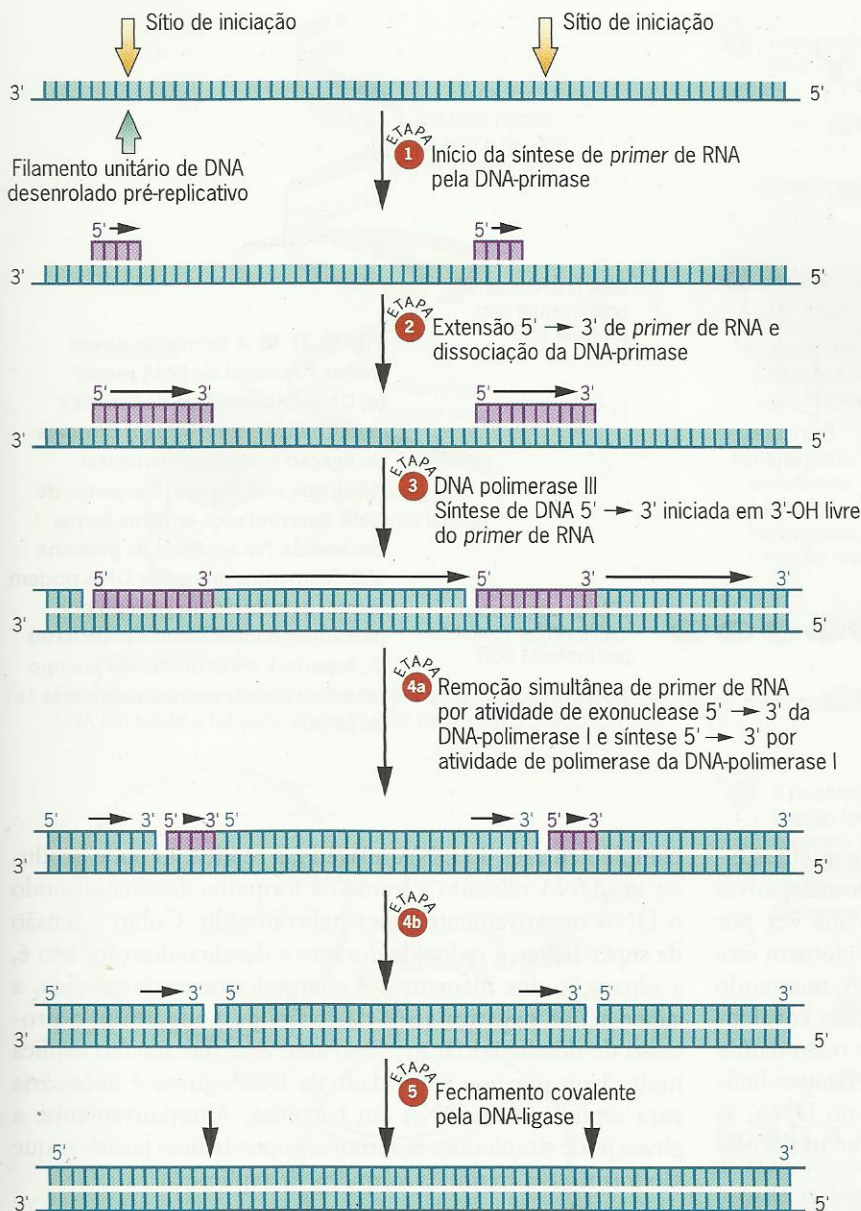


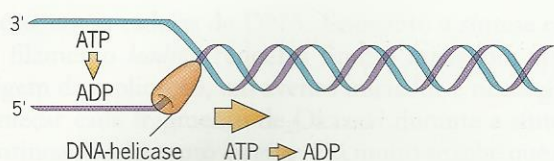
Fig. 10.20 ■ Síntese e substituição de *primers* de RNA durante replicação do filamento descontínuo do DNA. É sintetizado um curto filamento de RNA destinado a servir de 3'-OH *primer* para a síntese do DNA (veja Fig. 10.19). O *primer* de RNA é subsequentemente removido e substituído por DNA pelas atividades duplas de exonuclease 5' → 3' e polimerase 5' → 3' incluídas na DNA-polimerase I. A DNA-ligase então fecha covalentemente a cadeia de DNA nascente, catalisando a formação de ligações fosfodiéster entre fosfatos adjacentes 3'-hidroxila e 5'-fosfato (veja Fig. 10.18).

Sabe-se que tais estruturas em grampo impedem a atividade das DNA-polimerases. Em *E. coli*, a proteína SSB é codificada pelo gene *ssb*.

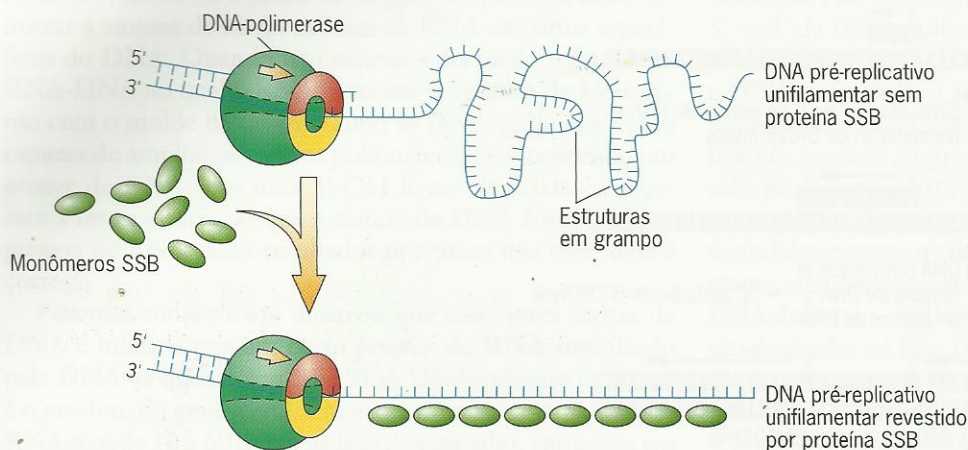
Lembre que o cromossomo de *E. coli* contém uma molécula circular de DNA. Com o giro do DNA de *E. coli* a 3.000 revoluções por minuto para permitir o desenrolamento dos filamentos parentais durante replicação (Fig. 10.22), o que serve de eixo de rotação que impede o DNA de se tornar mais helicoidizado (super-helicoidização positiva) à frente da forquilha de replicação? Os eixos de rotação necessários durante a replicação de moléculas circulares de DNA são fornecidos por enzimas chamadas **DNA-topoisomerases**. As topoisomerases catalisam quebras temporárias em moléculas de DNA, mas usam ligações covalentes para manter as moléculas clivadas. As topoisomerases são de dois tipos: (1) as enzimas DNA-topoisomerases I produzem quebras unifilamentares

temporárias ou cortes no DNA e (2) as enzimas DNA-topoisomerases II produzem quebras bifilamentares temporárias no DNA. Um resultado importante desta diferença é que as atividades da topoisomerase I removem super-hélices do DNA uma de cada vez, enquanto enzimas topoisomerases II removem e introduzem super-hélices duas de cada vez.

A quebra unifilar transitória produzida pela atividade da topoisomerase I fornece um eixo de rotação que permite que segmentos de DNA de lados opostos da quebra girem independentemente, com a ligação fosfodiéster no filamento intacto servindo como uma hélice (Fig. 10.23). As enzimas topoisomerases I são energeticamente eficientes. Elas conservam a energia das ligações fosfodiéster clivadas estocando-a em ligações covalentes entre si e os grupos fosfato nos sítios de clivagem. Elas então reaproveitam essa energia para ressoldar as quebras.



(a) DNA-helicase catalisa o desenrolamento da dupla hélice parental.



(b) Proteína de ligação ao DNA unifilamentar (SSB) mantém filamentos deselicoíizados em uma forma distendida para replicação.

Fig. 10.21 ■ A formação de um molde funcional de DNA requer (a) DNA-helicase, que desenrola a dupla hélice parental, e (b) proteína de ligação a DNA unifilamentar (SSB), que mantém os filamentos de DNA desenrolados em uma forma distendida. Na ausência da proteína SSB, filamentos únicos de DNA podem formar estruturas em grampo por pareamento intrafilamentar de bases (b, superior), e estruturas em grampo retardam ou interrompem a síntese de DNA.

As enzimas DNA-topoisomerases II induzem quebras bifilamentares transitórias e adicionam super-hélices negativas ou removem super-hélices positivas duas de cada vez por um mecanismo que requer energia (ATP). Elas efetuam este processo cortando ambos os filamentos do DNA, mantendo as pontas do sítio de clivagem mediante ligações covalentes, passando a dupla hélice intacta pelo corte e ressoldando a quebra (Fig. 10.24). Além de relaxar o DNA super-helicoidizado e introduzir super-hélices negativas no DNA, as enzimas DNA-topoisomerases II podem separar moléculas circulares entrelaçadas de DNA.

A topoisomerase tipo II mais bem caracterizada é uma enzima chamada **DNA-girase** em *E. coli*. A DNA-girase é um tetrâmero com duas subunidades α codificadas pelo gene *gyrA* (originalmente *nalA*, de ácido nalidíxico) e duas subunidades β especificadas pelo gene *gyrB* (antes *cou*, de coumermicina). O ácido nalidíxico e a coumermicina são antibióticos que bloqueiam a replicação do DNA em *E. coli* inibindo a atividade da DNA-girase. O ácido nalidíxico e a coumermicina inibem a síntese de DNA ligando-se às subunidades α e β , respectivamente, da DNA-girase. Assim, a atividade da DNA-girase é necessária para que ocorra replicação do DNA em *E. coli*.

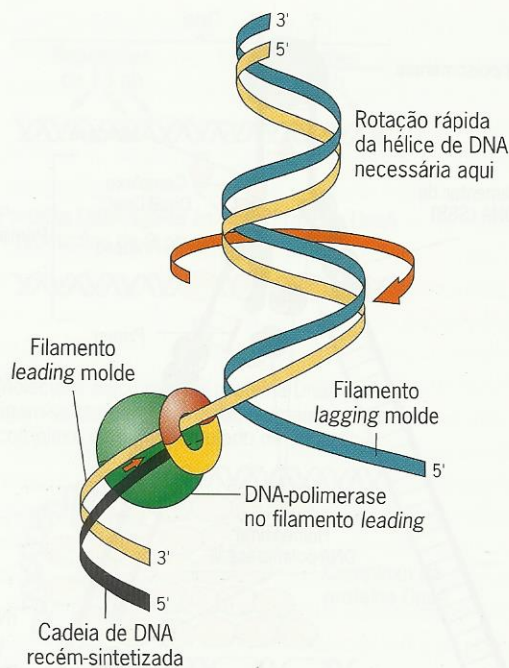
Lembre que o DNA cromossômico é negativamente super-helicoidizado em *E. coli* (Cap. 9). As super-hélices negativas em cromossomos bacterianos são introduzidas pela DNA-girase, com energia fornecida pelo ATP. Esta atividade da DNA-girase fornece outra solução para o problema de deselicoíização. Em vez de criar super-hélices positivas à frente da forquilha de replicação deselicoíizando os filamentos

complementares do DNA relaxado, a replicação pode produzir um DNA relaxado à frente da forquilha deselicoíizando o DNA negativamente super-helicoidizado. Como a tensão da super-hélice é reduzida durante a deselicoíização, isto é, a separação dos filamentos é energeticamente favorecida, a super-hélice negativa atrás da forquilha pode ativar o processo de deselicoíização. Caso sim, este mecanismo explica muito bem por que a atividade da DNA-girase é necessária para replicação do DNA em bactérias. Alternativamente, a girase pode simplesmente remover super-hélices positivas que são formadas à frente da forquilha de replicação.

O Aparelho de Replicação: Proteínas Pré-priming, Primossomos e Replissomos

As características básicas de replicação do DNA e a maioria dos componentes importantes do aparelho de replicação foram introduzidos nas seções anteriores deste capítulo. Agora juntaremos as peças e veremos as atividades coordenadas desses componentes durante a replicação do DNA cromossômico. Para fazer isto, consideraremos a seqüência de eventos que ocorrem durante a replicação da molécula circular do DNA no cromossomo de *E. coli*. A Fig. 10.25 mostra os componentes mais importantes do aparelho de replicação em *E. coli* e indica onde eles estão localizados em uma forquilha de replicação.

A replicação do cromossomo de *E. coli* começa em *oriC*, a seqüência única na qual é iniciada a replicação, com a formação de uma região localizada de separação do filamento cha-



(a) Para deslicoidizar os filamentos moldes em *E. coli*, a hélice de DNA em frente à forquilha de replicação deve girar a 3.000 rpm.



(b) Sem um giro ou eixo de rotação, o processo de desenrolamento produziria super-hélices positivas em frente às forquilhas de replicação.

Fig. 10.22 ■ Há necessidade de uma hélice ou eixo de rotação durante replicação de moléculas circulares de DNA como aquelas presentes nos cromossomos de *E. coli* ou no fago λ .

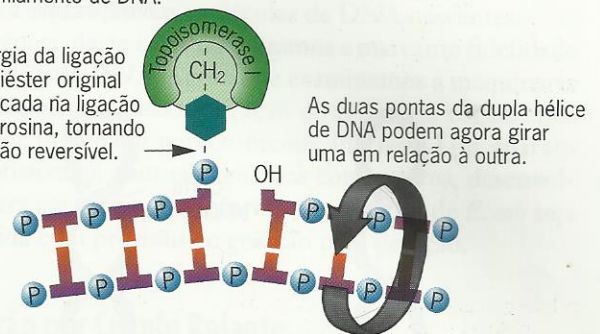
mada *bolha de replicação*. Essa bolha de replicação é formada pela interação de **proteínas pré-priming** com *oriC* (Fig. 10.26). A primeira etapa de pré-priming parece ser a ligação de quatro moléculas do produto do gene *dnaA*, a **proteína DnaA**, a quatro repetições com 9 pares de bases (pb) em *oriC*.

ETAPA 1 Uma ponta da dupla hélice de DNA não pode girar com relação à outra ponta.



ETAPA 2 A DNA-topoisomerase I liga-se covalentemente a um fosfato do DNA, quebrando assim uma ligação fosfodiéster em um filamento de DNA.

A energia da ligação fosfodiéster original é estocada na ligação fosfotirosina, tornando a reação reversível.



ETAPA 3 A reestruturação da ligação fosfodiéster regenera tanto a hélice de DNA quanto a DNA-topoisomerase de forma inalterada.



Fig. 10.23 ■ A DNA-topoisomerase I produz quebras unifilamentares temporárias no DNA que atuam como eixos de rotação, ou hélices, durante replicação do DNA.

Em seguida, as proteínas DnaA ligam-se cooperativamente para formar um cerne de 20 a 40 polipeptídeos com o DNA *oriC* enrolado na superfície do complexo protéico. A separação dos filamentos começa dentro de três repetições com 13 pb em tandem em *oriC* e se espalha até que seja criada a bolha de replicação. Um complexo da proteína DnaB (a DNA-helicase hexamérica) e a **proteína DnaC** (seis moléculas) juntam-se ao complexo de iniciação e contribuem para a formação de duas forquilhas bidirecionais de replicação. A proteína DnaT também está presente no complexo protéico pré-priming, mas sua função é desconhecida. Outras proteínas associadas ao complexo de iniciação em *oriC* são a proteína DnaJ, a proteína DnaK, a proteína PriA, a proteína PriB, a proteína PriC, a proteína HU de ligação ao DNA, a **DNA-girase** e a proteína unifilamentar de ligação ao DNA (SSB). Em alguns casos, entretanto, seu envolvimento fun-

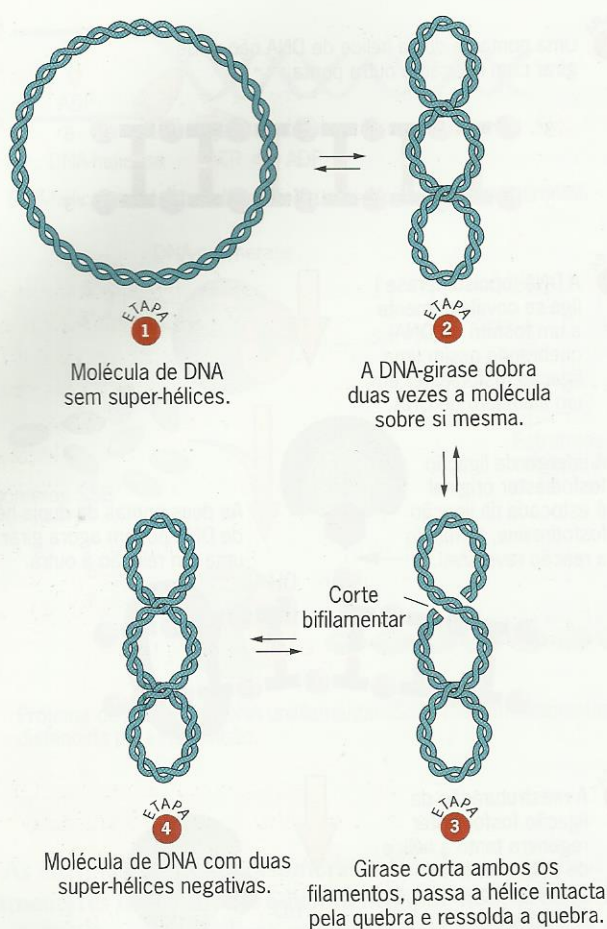


Fig. 10.24 ■ Mecanismo de ação da DNA-girase. Uma DNA-topoisomerase II de *E. coli* é necessária para replicação do DNA.

cional no processo pré-*priming* não foi estabelecido. Em outros casos, sabe-se que elas estão envolvidas, mas seus papéis são desconhecidos. A proteína DnaA parece ser em grande parte responsável pela separação localizada dos filamentos em *oriC* durante o processo de iniciação.

Uma vez formada a forquilha de replicação, a síntese de novos filamentos de DNA é iniciada pelos *primers* de RNA sintetizados pela DNA-primase. Um único *primer* de RNA é suficiente para a replicação contínua do filamento *leading*, mas a replicação descontínua do filamento *lagging* requer um *primer* de RNA para começar a síntese de cada fragmento de Okazaki. A iniciação dos fragmentos de Okazaki no filamento *lagging* é feita pelo **primossomo**, um complexo protéico contendo DNA-primase e DNA-helicase. O primossomo move-se ao longo de uma molécula de DNA, ativado pela energia do ATP. A DNA-helicase desenrola a dupla hélice e a DNA-primase sintetiza os *primers* de RNA para fragmentos sucessivos de Okazaki. Os *primers* de RNA são covalentemente ampliados com desoxirribonucleotídeos pela DNA-polimerase III. As DNA-topoisomerases fornecem quebras transitórias no DNA que servem como hélices para desenrolamento do DNA e para manter o DNA não

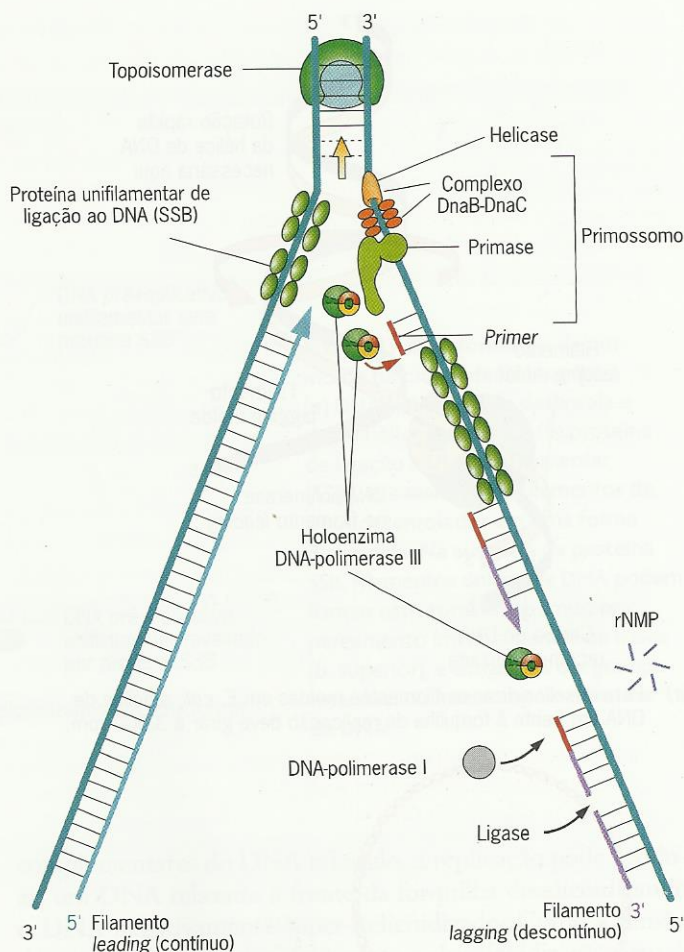


Fig. 10.25 ■ Diagrama de uma forquilha de replicação em *E. coli* mostrando os principais componentes do aparelho de replicação. rNMP = monofosfatos de ribonucleosídeos.

emaranhado. A proteína de ligação ao DNA unifilar reveste o DNA pré-replicative desenrolado e o mantém em um estado distendido para a DNA-polimerase III. Os *primers* de RNA são substituídos por DNA pela DNA-polimerase I, e as quebras unifilamentares deixadas pela polimerase são fechadas pela DNA-ligase. O DNA é então condensado no nucleóide, ou genoma dobrado, de *E. coli*, em parte por super-helicoidização negativa introduzida pela DNA-girase. Todas estas enzimas e as proteínas de ligação do DNA funcionam em conjunto em cada forquilha de replicação.

À medida que a forquilha de replicação se move ao longo de uma dupla hélice parental, os dois filamentos de DNA (o de replicação contínua e o de replicação descontínua) são replicados na série altamente coordenada de reações descritas anteriormente. O aparelho de replicação completo que se move ao longo da molécula de DNA em uma forquilha de replicação é chamado **replissomo** (Fig. 10.27). O replissomo contém a holoenzima DNA-polimerase III. Um cerne catalítico replica o filamento *leading*, o segundo cerne catalítico replica o filamento *lagging* e o primossomo desenrola

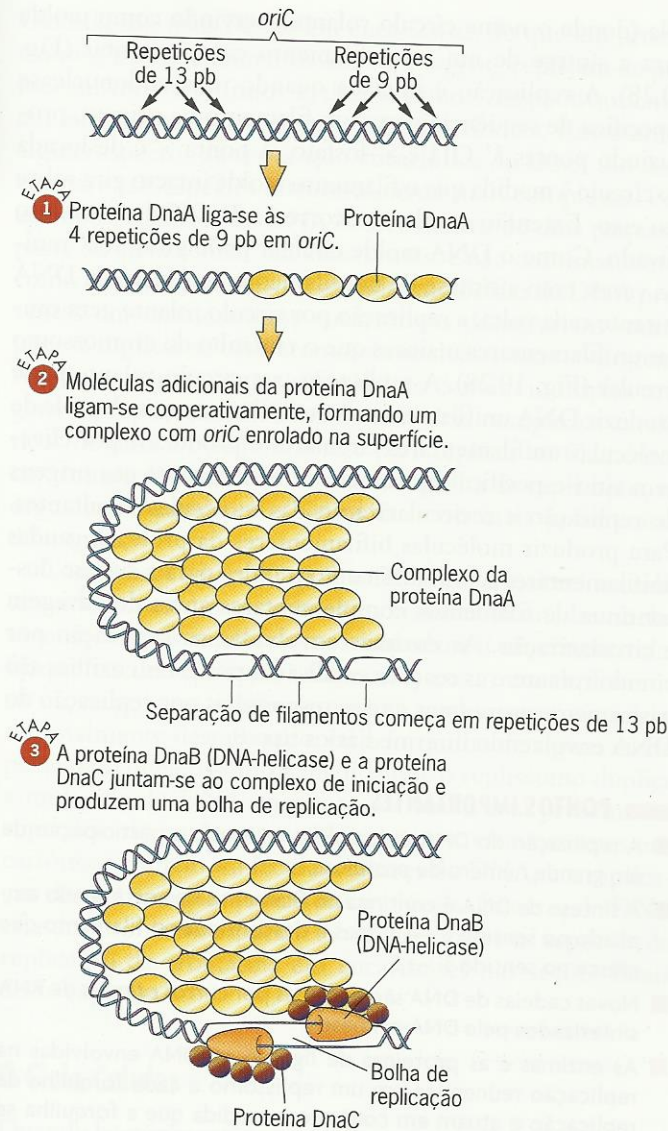


Fig. 10.26 ■ Pré-priming de replicação do DNA em *oriC* no cromossomo de *E. coli*.

a molécula de DNA parental e sintetiza os *primers* de RNA necessários para a síntese descontínua do filamento *lagging*. Para que os dois cernes catalíticos da holoenzima polimerase III sintetizem os filamentos *leading* e *lagging* nascentes, acredita-se que o filamento *lagging* forma uma alça do primossomo até o segundo cerne catalítico da DNA-polimerase III (Fig. 10.27).

Em *E. coli*, o término da replicação ocorre em sítios variáveis dentro de regiões chamadas *terA* e *terB*, o que bloqueia o movimento de forquilha de replicação avançando nos sentidos anti-horário e horário, respectivamente. As DNA-topoisomerases ou combinações especiais de enzimas então facilitam a separação das moléculas de DNA nascentes.

No começo deste capítulo, notamos a marcante fidelidade da replicação do DNA. Agora que examinamos a maquinaria celular responsável pela replicação do DNA em organismos vivos, essa fidelidade parece menos incrível. Um aparato muito sofisticado, com salvaguardas contra erros, desenvolveu-se para garantir que a informação genética de *E. coli* seja transmitida com precisão de geração para geração.

Replicação por Círculo Rolante

Nas seções anteriores deste capítulo, consideramos os DNA replicativos nas formas θ , de olho e de Y. Examinaremos agora outro tipo importante de replicação do DNA chamado **replicação por círculo rolante**. A replicação por círculo rolante é usada (1) por muitos vírus para duplicar seus genomas, (2) em bactérias para transferir o DNA de células doadoras para receptoras durante um tipo de troca genética (Cap. 8) e (3) em anfíbios para amplificar DNA extracromossômico que possuem grupos de genes de RNA ribossômico durante ovocitogênese (Cap. 21).

Como o nome indica, a replicação por círculo rolante é um mecanismo para replicar moléculas circulares de DNA. O aspecto único da replicação em círculo rolante é que um filamento de DNA circular parental permanece intacto e

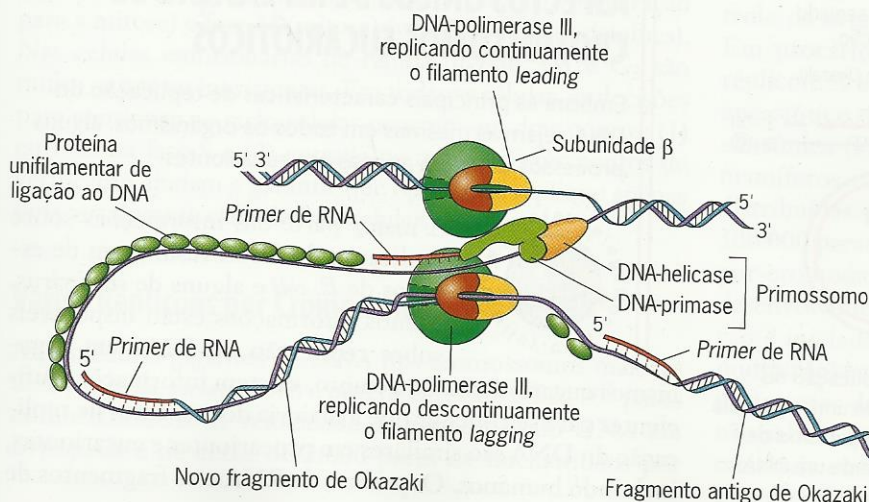


Fig. 10.27 ■ Diagrama do replissomo de *E. coli* mostrando os dois cernes catalíticos da DNA-polimerase III replicando os filamentos contínuo e descontínuo e o primossomo desenrolando a dupla hélice parental e iniciando a síntese de novas cadeias com *primers* de RNA. Todo o replissomo move-se ao longo da dupla hélice parental, com cada componente fazendo sua respectiva função de um modo conjunto. De fato, o complexo de replicação provavelmente não se move. Em vez disso, o DNA é levado através do replissomo. A replicação ocorre da esquerda para a direita.

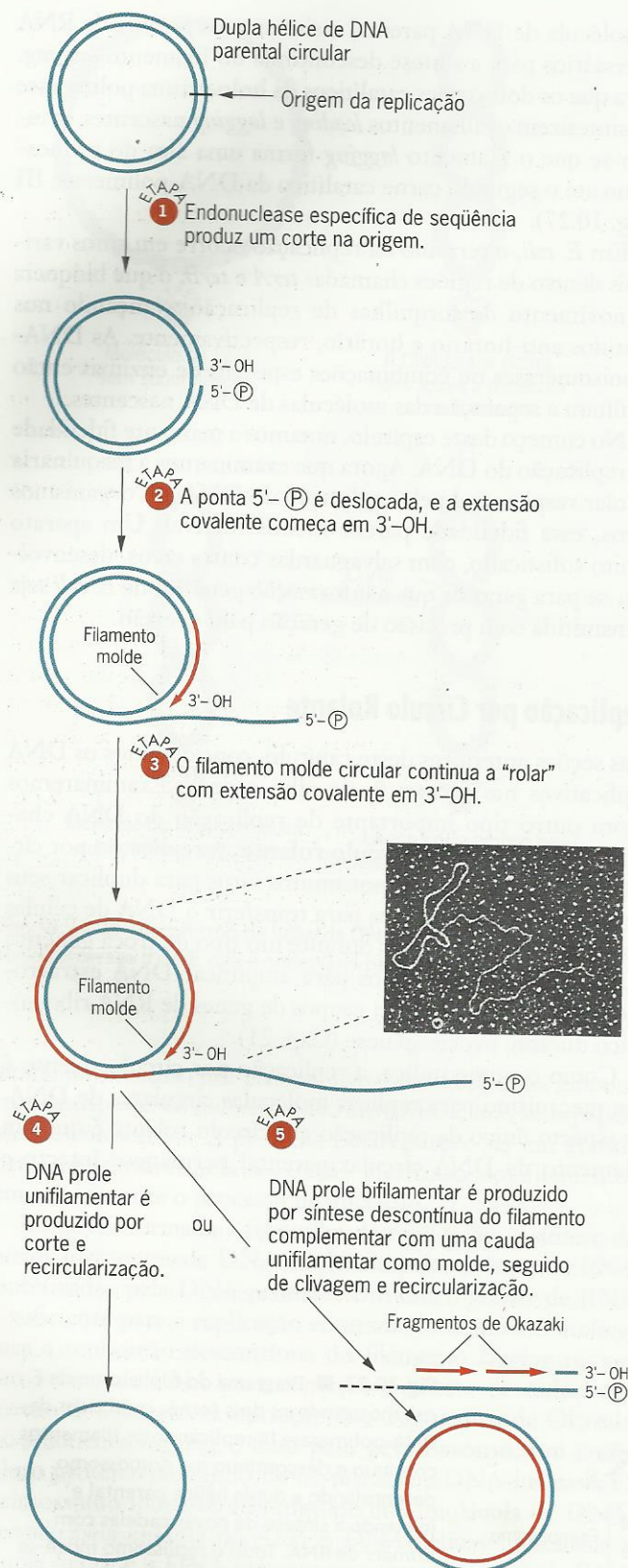


Fig. 10.28 ■ O mecanismo de círculo rolante de replicação do DNA. O detalhe mostra uma micrografia eletrônica de uma molécula de DNA do bacteriófago ϕ X174 replicando-se pelo mecanismo de círculo rolante. Uma cauda unifilamentar prolonga-se de um DNA replicativo, circular, bifilamentar.

rola (donde o nome círculo rolante), servindo como molde para a síntese de um novo filamento complementar (Fig. 10.28). A replicação é iniciada quando uma endonuclease específica de sequência corta um filamento na origem, produzindo pontas 3'-OH e 5'-fosfato. A ponta 5' é deslocada do círculo à medida que o filamento molde intacto gira sobre seu eixo. Extensão covalente ocorre na 3'-OH do filamento clivado. Como o DNA molde circular pode girar 360° muitas vezes, com a síntese de um filamento completo de DNA durante cada volta, a replicação por círculo rolante gera caudas unifilamentares maiores que o tamanho do cromossomo circular (Fig. 10.28). A replicação por círculo rolante pode produzir DNA unifilamentares ou bifilamentares. A prole de moléculas unifilamentares circulares é produzida por clivagem sítio-específica das caudas unifilamentares nas origens de replicação e recircularização das moléculas resultantes. Para produzir moléculas bifilamentares da prole, as caudas unifilamentares são usadas como moldes para a síntese descontínua de filamentos complementares antes de clivagem e circularização. As enzimas envolvidas na replicação por círculo rolante e as reações catalisadas por estas enzimas são basicamente as mesmas que as responsáveis por replicação do DNA envolvendo intermediários tipo θ .

PONTOS IMPORTANTES

- A replicação do DNA é complexa, exigindo a participação de um grande número de proteínas.
- A síntese de DNA é contínua no filamento que está sendo ampliado no sentido $5' \rightarrow 3'$, mas é descontínua no filamento que cresce no sentido $3' \rightarrow 5'$.
- Novas cadeias de DNA são iniciadas por curtos *primers* de RNA sintetizados pela DNA-primase.
- As enzimas e as proteínas de ligação ao DNA envolvidas na replicação reúnem-se em um replissomo a cada forquilha de replicação e atuam em conjunto à medida que a forquilha se move ao longo da molécula de DNA parental.

ASPECTOS ÚNICOS DE REPLICAÇÃO DE CROMOSSOMOS EUCARIÓTICOS

Embora as principais características de replicação do DNA sejam as mesmas em todos os organismos, alguns processos ocorrem apenas em eucariontes.



A maior parte das informações sobre replicação do DNA resultaram de estudos de *E. coli* e alguns de seus vírus. Menos informações estão disponíveis sobre replicação do DNA em organismos eucarióticos. Entretanto, existem informações suficientes para se concluir que a maioria dos aspectos de replicação do DNA são similares em procariontes e eucariontes, incluindo humanos. Os *primers* de RNA e os fragmentos de

Okazaki são mais curtos em eucariontes do que em procariontes, mas os filamentos *leading* e *lagging* replicam-se por mecanismos contínuos e descontínuos, respectivamente, em eucariontes tanto quanto em procariontes. Entretanto, alguns aspectos da replicação eucariótica do DNA são únicos para tais espécies estruturalmente mais complexas. Por exemplo, a síntese de DNA ocorre dentro de uma pequena parte do ciclo celular em eucariontes, e não continuamente como nos procariontes. As grandes moléculas de DNA presentes em cromossomos eucarióticos demorariam muito para se replicar se cada cromossomo tivesse uma só origem. Assim, os cromossomos eucarióticos contêm múltiplas origens de replicação. Em vez de usar dois complexos catalíticos de uma DNA-polimerase para replicar os filamentos contínuo e descontínuo de cada forquilha de replicação, os organismos eucarióticos usam duas ou mais polimerases diferentes.

Como discutido no Cap. 9, o DNA eucariótico é acondicionado em estruturas contendo histonas chamadas nucleossomos. Esses nucleossomos impedem o movimento das forquilhas de replicação? Caso não, como um replissomo se move por um nucleossomo? O nucleossomo é parcial ou totalmente desmontado, ou a forquilha de algum modo passa pelo nucleossomo à medida que o replissomo duplica a molécula de DNA enquanto ela ainda está presente na superfície do nucleossomo? Finalmente, cromossomos eucarióticos contêm moléculas lineares de DNA, e a replicação descontínua das pontas das moléculas lineares de DNA cria um problema especial. Abordaremos estes aspectos de replicação da cromatina em eucariontes nas seções finais deste capítulo.

O Ciclo Celular

Quando bactérias estão crescendo em meio rico, a replicação do DNA ocorre sem parar durante o ciclo celular. Entretanto, em eucariontes, a replicação do DNA é restrita à fase S (de síntese; Cap. 2). Lembre que o ciclo celular eucariótico normal consiste em fase G₁ (imediatamente após o término da mitose; G de *gap*, intervalo), fase S, fase G₂ (preparação para a mitose) e fase M (mitose) (veja Cap. 2 para detalhes). Nas células embrionárias de rápida divisão, G₁ e G₂ são muito curtas ou inexistentes. Em todas as células, as decisões para continuar no ciclo celular ocorrem em dois pontos: (1) entrada na fase S e (2) entrada na mitose. Estes pontos de verificação ajudam a garantir que o DNA se replique apenas uma vez durante cada divisão celular.

Vários Réplicons por Cromossomo

As moléculas gigantes do DNA nos cromossomos maiores de *Drosophila melanogaster* contêm cerca de $6,5 \times 10^7$ pares de nucleotídeos. A velocidade de replicação do DNA em *Drosophila* é de cerca de 2.600 pares de nucleotídeos por

minuto a 25°C. Uma única forquilha de replicação levaria, portanto, 17,5 dias para replicar uma dessas moléculas gigantes de DNA. Com duas forquilhas de replicação movendo-se bidirecionalmente de uma origem central, tal molécula de DNA seria replicada em apenas 8,5 dias. Considerando que os cromossomos dos embriões de *Drosophila* replicam-se dentro de 3 a 4 minutos e os núcleos dividem-se uma vez a cada 9 a 10 minutos durante as primeiras divisões de clivagem, fica claro que cada molécula gigante de DNA deve conter muitas origens de replicação. De fato, a replicação completa do DNA do maior cromossomo de *Drosophila* dentro de 3,5 minutos precisaria de 7.000 forquilhas de replicação distribuídas em intervalos iguais ao longo da molécula. Assim, são necessárias várias origens de replicação para permitir que moléculas muito grandes de DNA em cromossomos eucarióticos repliquem-se dentro dos tempos observados de divisão celular.

A primeira evidência de várias origens em cromossomos eucarióticos resultou de experimentos de marcação de pulso com células de hamster chinês crescendo em cultura. Em 1968, quando Joel Huberman e Arthur Riggs fizeram marcação de pulsos em células com ³H-timidina por alguns minutos, extraída do DNA, e fizeram análises auto-radiográficas do DNA marcado, eles observaram arranjos em tandem de grãos de prata expostos (Fig. 10.29a). A interpretação mais simples de seus resultados é que macromoléculas individuais de DNA contêm várias origens de replicação. Quando o período de marcação de pulso foi seguido de um curto intervalo de crescimento em meio não radioativo (experimentos de pulso-caça), os arranjos em tandem continham regiões centrais de alta densidade de grãos com caudas de densidade diminuída de grãos em ambas as pontas (Fig. 10.29b). Este resultado indica que a replicação em eucariontes é bidirecional, como o é na maioria dos procariontes. As caudas de densidades diminuídas de grãos resultam da diluição gradual das concentrações intracelulares de ³H-timidina por ¹H-timidina, à medida que as forquilhas de replicação se movem bidirecionalmente das origens centrais para os terminais de replicação (Fig. 10.29c).

Um segmento de DNA cuja replicação está sob o controle da origem e de dois terminais é chamado **réplicon**. Em procariontes, todo o cromossomo geralmente é um replicon. A existência de vários replicons por cromossomo eucariótico foi diretamente verificada por microscopia eletrônica (Fig. 10.30). Os genomas de humanos e outros mamíferos contêm cerca de 10.000 origens de replicação distribuídas pelos cromossomos a intervalos de 30.000 a 300.000 pares de bases. Claramente, o número de replicons por cromossomo não é fixado durante o crescimento e o desenvolvimento de um eucarionte multicelular. A replicação é iniciada em mais sítios durante as divisões celulares muito rápidas da embriogênese do que durante os estágios posteriores do desenvolvimento. Infelizmente, geneticistas não sabem que fatores determinam quais origens são operacionais em determinado momento ou em determinado tipo de célula.

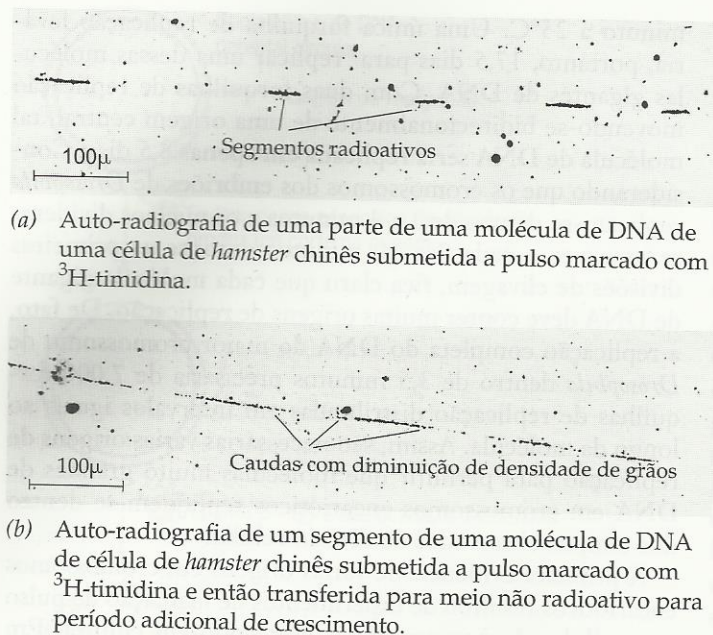
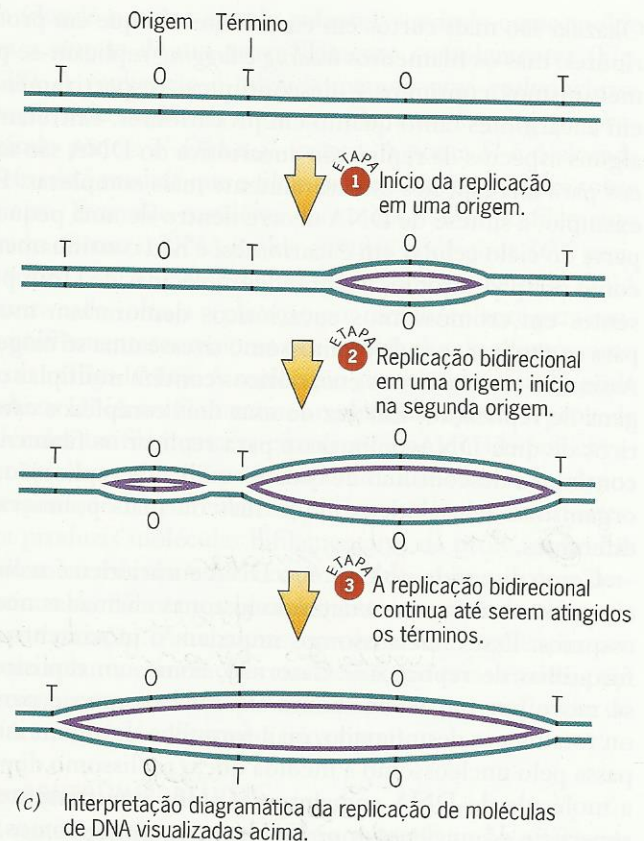


Fig. 10.29 ■ Evidência de replicação bidirecional de vários réplicons em moléculas gigantes de DNA de eucariontes. As disposições em tandem de radioatividade em (a) indicam que a replicação ocorre em várias origens; as caudas com densidades decrescentes de grãos observadas em (b) indicam que a replicação ocorre bidirecionalmente a partir de cada origem (c).



Duas ou Mais DNA-polimerases em uma Única Forquilha de Replicação

Em virtude da complexidade do replissomo na bactéria simples *E. coli* (veja Figs. 10.25 e 10.27), parece provável que o aparelho de replicação é ainda mais complexo em eucariontes. Embora o conhecimento da estrutura da maquinaria replicativa em eucariontes ainda seja limitado, muitas características de replicação do DNA são similares em eucariontes e procariontes.

Como no caso dos procariontes, grande parte da informação sobre a síntese de DNA em eucariontes veio do desenvolvimento e da dissecação dos sistemas *in vitro* de replicação do DNA. Estudos da replicação de vírus com DNA de eucariontes demonstraram-se informativos, e, desses vírus, o vírus 40 de símios (SV40) foi particularmente útil. A replicação do SV40 é feita quase que totalmente pelo aparelho de replicação da célula hospedeira. Apenas uma proteína viral, o chamado antígeno T, é necessária para replicação do cromossomo de SV40.

Como em procariontes, a deslicoidização dos filamentos de DNA parental requer uma DNA-topoisomerase e uma DNA-helicase. Os filamentos deslicoidizados são mantidos no estado distendido por uma proteína de ligação ao DNA unifilar chamada proteína A de replicação (Rp-A). Entretanto, ao contrário dos processos que ocorrem em procariontes, a replicação do DNA cromossômico em eucariontes requer a atividade de três DNA-polimerases diferentes – a polimerase α (Pol α), a polimerase δ (Pol δ) e a polimerase ϵ

(Pol ϵ). Pelo menos duas polimerases, talvez todas as três, estão presentes em cada forquilha de replicação (replissomo), e cada polimerase contém várias subunidades. Também, enquanto o replissomo de *E. coli* contém 13 proteínas conhecidas, os replissomos de leveduras e mamíferos contêm pelo menos 27 componentes diferentes.

Em eucariontes, Pol α é necessária para o início da replicação nas origens e para inserir o *primer* de fragmentos de Okazaki durante a síntese descontínua do filamento *lagging*. Pol α existe em um complexo estável com a DNA-primase. De fato, elas se purificam durante o isolamento. A primase sintetiza os *primers* de RNA, que são então ampliados com desoxirribonucleotídeos pela Pol α para produzir uma cadeia de RNA-DNA com cerca de 30 nucleotídeos de tamanho total. Essas cadeias *primer* de RNA-DNA são então ampliadas por Pol δ (talvez Pol ϵ em alguns casos). Acredita-se que a Pol δ catalise a maior parte da síntese processiva do DNA cromossômico. A Pol δ deve interagir com proteínas PCNA (antígeno nuclear de proliferação celular) e o fator C de replicação (Rf-C) para ser ativa (Fig. 10.31). PCNA é o grampo deslizante que prende a Pol δ ao DNA para permitir replicação processiva (com o objetivo de evitar que a polimerase saia do molde). PCNA é equivalente à subunidade β da DNA-polimerase III em *E. coli* (veja Fig. 10.14). Rf-C é necessária para que PCNA carregue o DNA. PCNA é uma proteína trimérica que forma um anel fechado; Rf-C induz uma mudança de conforma-

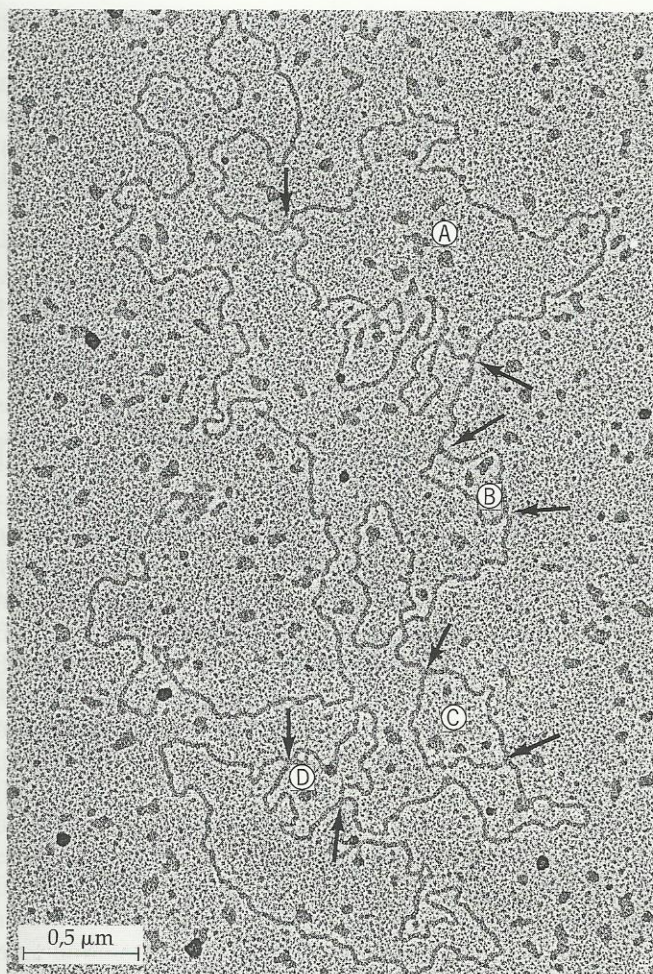


Fig. 10.30 ■ Micrografia eletrônica de uma molécula de DNA em *D. melanogaster* mostrando múltiplos sítios de replicação. São apresentadas quatro estruturas de replicação em forma de olho (A-D) (marcadas de A-D) no segmento da molécula de DNA mostrada. As setas mostram as posições das forquilhas de replicação.

ção de PCNA que lhe permite circular o DNA, fornecendo o grampo deslizante essencial.

As polimerases δ e ϵ contêm a atividade de exonuclease $3' \rightarrow 5'$ necessária para revisão (veja Fig. 10.16). Entretanto, elas não têm atividade de exonuclease $5' \rightarrow 3'$. Assim, elas não podem remover *primers* de RNA como a DNA-polimerase I de *E. coli* faz. Em vez disso, os *primers* de RNA são removidos por duas nucleases, a ribonuclease H1 (que degrada o RNA presente nos dúplexes RNA-DNA) e a ribonuclease FEN-1 (F1-nuclease 1). A Pol δ então preenche os espaços, e a DNA-ligase fecha os cortes, produzindo filamentos da prole ligados covalentemente.

A DNA-polimerase ϵ é necessária para replicação *in vivo* e para reparo de alguns tipos de danos ao DNA. Entretanto, sua(s) função(ões) exata(s) é(são) incerta(s). Uma proposta é que a Pol δ seja responsável pela síntese tanto do filamento contínuo quanto do descontínuo, tendo a Pol ϵ funções especializadas desconhecidas. Outro modelo sugere que a Pol δ replica o filamento descontínuo e a Pol ϵ replica o filamento contínuo. Um terceiro modelo tem seus papéis revertidos. Claramente, são necessárias mais pesquisas para determinar as funções das múltiplas polimerases e de outras proteínas no replissomo eucariótico.

Duplicação de Nucleossomos nas Forquilhas de Replicação

Como discutimos no Cap. 9, o DNA nos cromossomos interfásicos eucarióticos é acondicionado em contas com aproximadamente 10 nm chamadas nucleossomos. Cada nucleossomo contém 166 pares de nucleotídeos de DNA enrolados em duas voltas de um octâmero de moléculas de histonas. Em vista do tamanho dos nucleossomos e do grande tamanho dos replissomos de DNA, parece improvável que uma forquilha de replicação possa mover-se por um nucleossomo

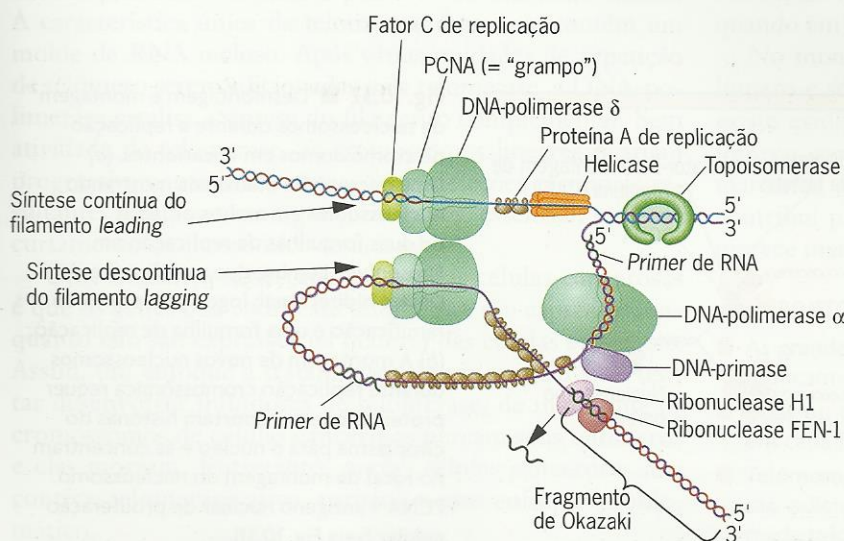


Fig. 10.31 ■ Alguns componentes importantes de um replissomo em eucariotes. Cada replissomo contém duas polimerases diferentes, α e δ (ou talvez ϵ). O complexo DNA-polimerase α -DNA-primase sintetiza os *primers* de RNA e adiciona segmentos curtos de DNA. A DNA-polimerase δ (às vezes ϵ) então completa a síntese dos fragmentos de Okazaki no filamento descontínuo e catalisa a síntese contínua no filamento *leading*. PCNA (antígeno nuclear de proliferação celular) é equivalente à subunidade β da DNA-polimerase III de *E. coli*. Ele liga a polimerase δ à molécula de DNA facilitando a síntese de longas cadeias de DNA. As ribonucleases H1 e FEN-1 (nuclease 1 F1) remove os *primers* de RNA, a polimerase δ preenche os espaços e a DNA-ligase (não mostrada) fecha os cortes, como na *E. coli* (veja Fig. 10.20).

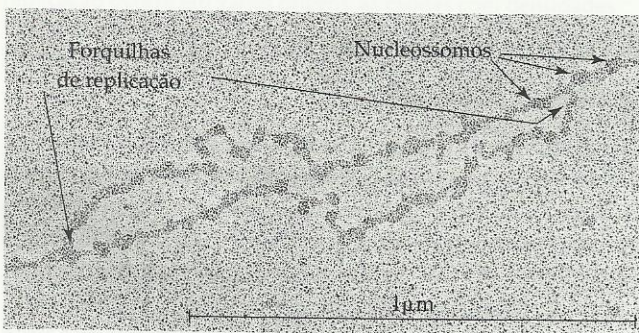
intacto. As micrografias eletrônicas de cromatina replicante em *Drosophila* mostram claramente nucleossomos com estrutura aproximadamente normal e espaçamento em ambos os lados das forquilhas de replicação (Fig. 10.32a); isto é, os nucleossomos parecem ter a mesma estrutura e espaçamento antes de uma forquilha de replicação (DNA pós-replicativo) que têm na frente de uma forquilha de replicação (DNA pré-replicativo). Assim, vários modelos foram propostos para alterações transitórias na estrutura do nucleossomo que permitiriam que os replissomos sintetizem DNA à medida que passam pelos nucleossomos. Em um modelo, o nucleossomo divide-se em duas metades de nucleossomos enquanto o replissomo passa, e as metades então se reúnem em nucleossomos intactos atrás da forquilha de replicação.

Independente do mecanismo pelo qual a forquilha de replicação passa pelos nucleossomos, a replicação do DNA e a montagem do nucleossomo são bem acopladas em eucariontes. Como a massa das histonas em nucleossomos é equivalente à do DNA, grandes quantidades de histonas devem ser sintetizadas durante cada geração celular para que os nucleossomos se dupliquem. Embora a síntese de histonas ocorra ao longo do ciclo celular, há aumento da biossíntese de histonas durante a fase S que gera histonas suficientes para duplicação da cromatina.

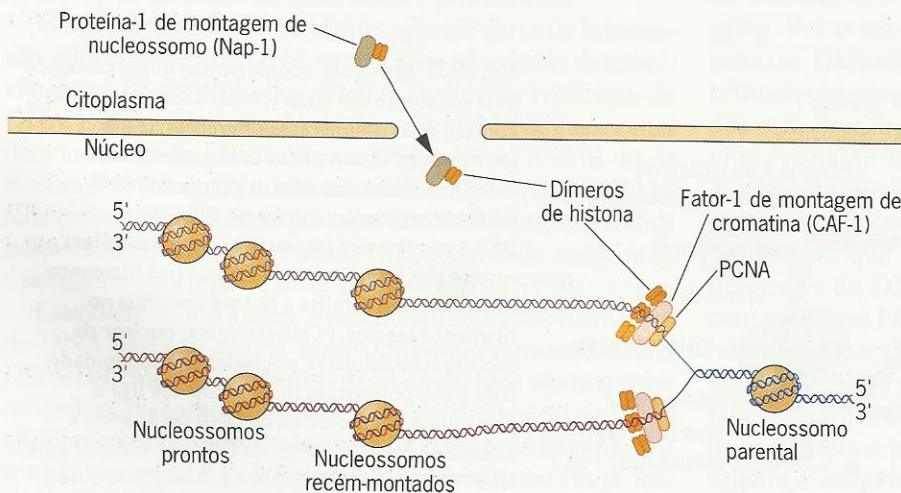
Quando experimentos de transferência de densidade foram feitos para examinar o modo de duplicação do nucleossomo, observou-se que os nucleossomos de ambas as moléculas da prole de DNA continham tanto complexos de histona antigos (pré-replicativos) quanto complexos novos (pós-replicativos). Assim, em nível de proteínas, a duplicação de nucleossomos parece ocorrer por um mecanismo dispersivo.

Várias proteínas estão envolvidas na desmontagem e montagem de nucleossomos durante a replicação cromossômica em eucariontes. Duas das mais importantes são a proteína 1 de montagem de nucleossomos (Nap-1) e o fator 1 de montagem da cromatina (CAF-1). Nap-1 transporta histonas de seu local de síntese no citoplasma para o núcleo e CAF-1 as leva para os locais cromossômicos de montagem do nucleossomo (Fig. 10.32b). CAF-1 leva histonas para os sítios de replicação do DNA ligando PNCA (antígeno nuclear de proliferação celular) – o grampo que prende a DNA-polimerase δ ao molde de DNA (veja Fig. 10.31). CAF-1 é uma proteína essencial em *Drosophila*, mas não em leveduras em que outras proteínas desempenham algumas de suas funções.

Muitas outras proteínas afetam a estrutura do nucleossomo. Algumas estão envolvidas em remodelamento da cromatina, mudando a estrutura do nucleossomo de modos



(a) Micrografia eletrônica mostrando nucleossomos em ambos os lados de cada uma das duas forquilhas de replicação em *Drosophila*.



(b) Alguns componentes necessários para montagem do nucleossomo durante replicação cromossômica em eucariontes.

Fig. 10.32 ■ Desmontagem e montagem de nucleossomos durante a replicação de cromossomos em eucariontes. (a) Uma micrografia eletrônica mostrando nucleossomos em ambos os lados de duas forquilhas de replicação em *Drosophila*. Lembre que a replicação do DNA é bidirecional; logo, cada ponto de ramificação é uma forquilha de replicação. (b) A montagem de novos nucleossomos durante replicação cromossômica requer proteínas que transportam histonas do citoplasma para o núcleo e as concentram no local de montagem do nucleossomo. PCNA = antígeno nuclear de proliferação celular (veja Fig. 10.31).

que ativam ou silenciam a expressão dos genes ali contidos. Outras modificam a estrutura do nucleossomo adicionando grupos metil ou acetil a histonas específicas. Além disso, eucariontes contêm várias histonas com estruturas um pouco diferentes das histonas principais, e a incorporação dessas histonas secundárias a nucleossomos pode mudar sua estrutura. Em *Drosophila*, por exemplo, a incorporação da histona H3.3 a nucleossomos resulta em níveis altos de transcrição dos genes ali existentes. Assim, a estrutura do nucleossomo não é invariante. Ao contrário, ela desempenha papel importante em modular a expressão gênica (veja Cap. 21).

Telomerase: Replicação das Pontas do Cromossomo

Discutimos as estruturas únicas de telômeros, ou pontas cromossômicas, no Cap. 9. Um motivo inicial para pensar que telômeros devem ter estruturas especiais foi que as DNA-polimerases não podem replicar o segmento terminal do DNA do filamento descontínuo de um cromossomo linear. Na ponta da molécula de DNA que se replica descontinuamente, não haveria filamento de DNA para dar uma 3'-OH livre (*primer*) para polimerização de desoxirribonucleotídeos após o *primer* de RNA do fragmento de Okazaki terminal ter sido removido (Fig. 10.33a). Ou (1) o telômero tem de ter uma estrutura única que facilita sua replicação ou (2) deve haver uma enzima especial que resolva este enigma de replicação da ponta do filamento descontínuo. De fato, evidências mostraram que ambos estão corretos. A estrutura especial de telômeros fornece um mecanismo simples para a adição de telômeros por uma enzima chamada **telomerase** contendo RNA.

Os telômeros de humanos, que contêm a sequência TTAGGG repetida em tandem, serão usados para ilustrar a maneira como a telomerase adiciona pontas a cromossomos (Fig. 10.33b). A telomerase reconhece a sequência do telômero rica em G na ponta 3' que se projeta e a amplia de 5' → 3' uma unidade de repetição a cada vez. A telomerase não preenche o espaço oposto à ponta 3' do filamento molde. Ela simplesmente estende a ponta 3' do filamento molde. A característica única da telomerase é que ela contém um molde de RNA incluso. Após várias unidades de repetição de telômero serem adicionadas pela telomerase, a DNA-polimerase catalisa a síntese do filamento complementar. Sem atividade de telomerase, os cromossomos lineares ficariam progressivamente mais curtos. Se as deleções terminais resultantes atingissem um gene, ou genes, essencial, este encurtamento cromossômico seria letal.

Uma mudança observada em muitas células cancerosas é que os genes que codificam telomerase são expressos, enquanto não são expressos na maioria das células somáticas. Assim, um enfoque a tratamentos de câncer tem sido tentar desenvolver inibidores de telomerase, de modo que os cromossomos de células cancerosas percam seus telômeros e elas morram. Entretanto, outras células cancerosas não contêm telomerase ativa, tornando este enfoque problemático.

Tamanho do Telômero e Envelhecimento em Humanos

Ao contrário das células da linhagem germinativa, a maioria das células somáticas humanas não tem atividade de telomerase. Quando células somáticas humanas são cultivadas, elas se dividem por um número limitado de vezes (geralmente apenas 20 a 70 gerações de células) antes que ocorra senescência e morte. Quando o tamanho dos telômeros é medido em várias culturas de células somáticas, é observada uma correlação entre o tamanho do telômero e o número de divisões celulares que precedem a senescência e a morte. Células com telômeros longos sobrevivem mais – passam por mais divisões celulares – do que células com telômeros mais curtos. Como seria esperado na ausência de atividade de telomerase, o tamanho do telômero diminui à medida que a idade da cultura de células aumenta. Ocasionalmente, observa-se que células somáticas adquirem a capacidade de proliferar indefinidamente em cultura, e foi demonstrado que tais células imortais têm atividade de telomerase, ao contrário de suas genitoras. Como uma das características comuns de todos os cânceres é a divisão celular descontrolada, ou imortalidade, cientistas propuseram que um modo de combater cânceres humanos seria inibir a atividade de telomerase em células cancerosas.

Outra evidência de uma correlação entre tamanho do telômero e envelhecimento em humanos veio de estudos de pessoas com distúrbios chamados **progerias**, doenças hereditárias caracterizadas por envelhecimento prematuro. Na forma mais grave de progeria, a síndrome Hutchinson-Gilford (Fig. 10.34), a senescência – rugas, calvície e outros sintomas de envelhecimento – começa imediatamente após o nascimento, e a morte em geral ocorre na adolescência. Em uma forma menos grave, a síndrome Werner, a senescência começa na puberdade, ocorrendo morte na faixa dos 40 anos. Consistente com a hipótese de que a diminuição do telômero contribui para o processo de envelhecimento, as células somáticas de pessoas com progeria têm telômeros curtos e apresentam capacidade proliferativa diminuída quando em meio de cultura.

No momento, a correspondência entre tamanho do telômero e senescência celular é totalmente correlativa. Não existe evidência direta indicando que encurtamento do telômero causa envelhecimento. Entretanto, a correlação é marcante, e a hipótese de que o encurtamento do telômero contribui para o processo de envelhecimento em humanos merece mais estudos.

PONTOS IMPORTANTES

- As grandes moléculas de DNA em cromossomos eucarióticos replicam-se bidirecionalmente a partir de várias origens.
- Duas ou três DNA-polimerases (α , δ e/ou ϵ) estão presentes em cada forquilha de replicação em eucariontes.
- Telômeros, as sequências únicas nas pontas dos cromossomos, são adicionados a cromossomos por uma enzima única chamada telomerase.

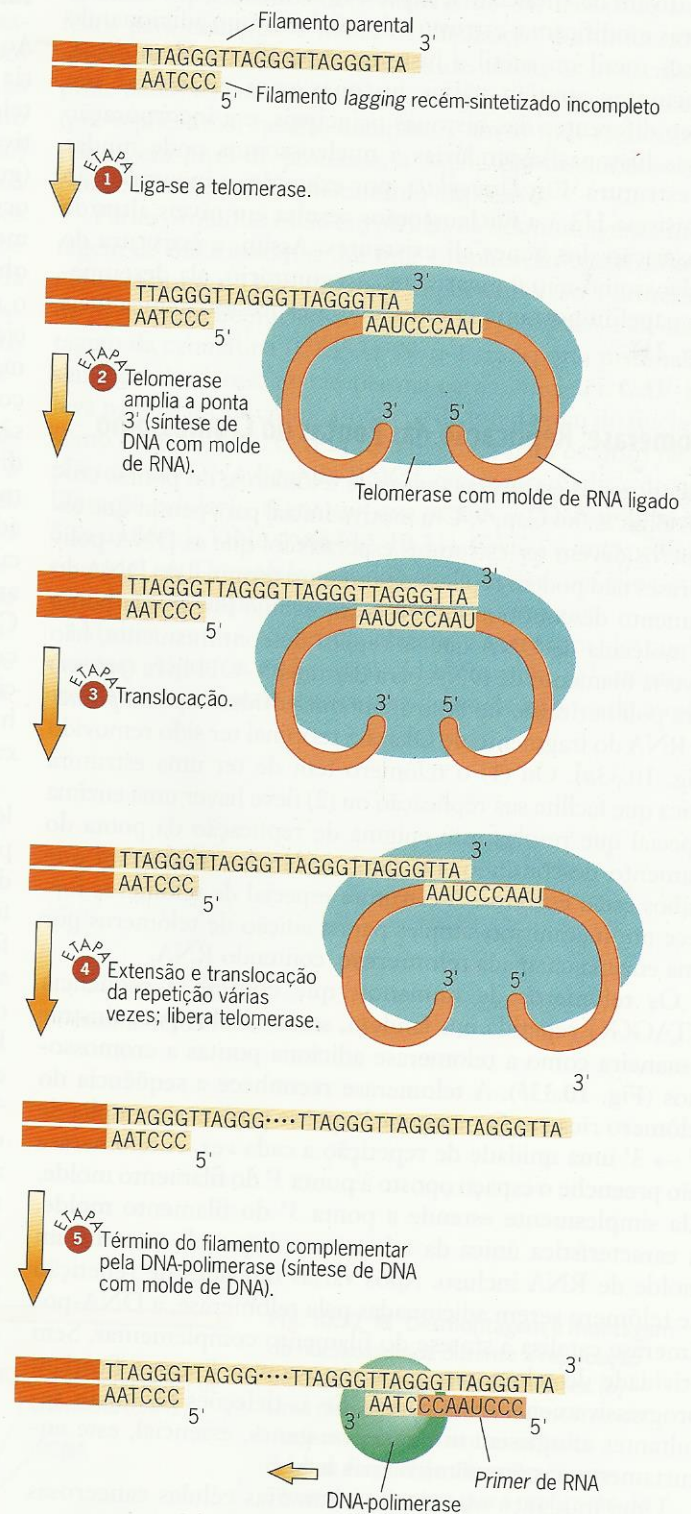
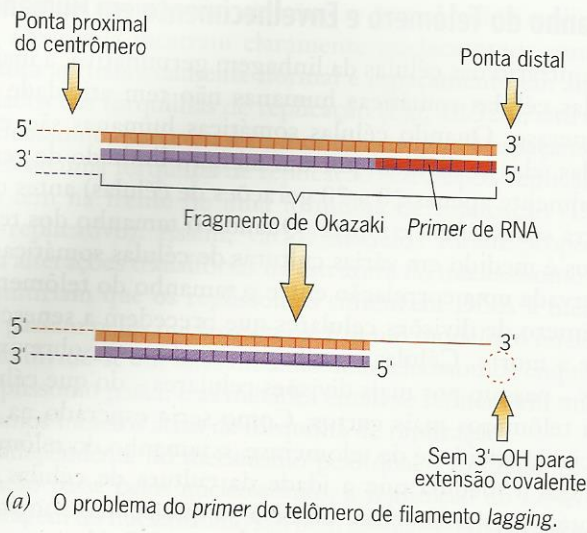


Fig. 10.33 ■ Replicação de telômeros cromossômicos. (a) Devido à necessidade de uma 3'-OH livre no final do filamento *primer*, as DNA-polimerases não podem substituir um *primer* de RNA que inicie a síntese de DNA perto ou na ponta do filamento *lagging*. (b) Estas pontas de cromossomos são replicadas por uma enzima especial chamada telomerase, que impede que os cromossomos fiquem mais curtos durante cada replicação. A sequência de nucleotídeos ao final do filamento descontínuo é especificada por uma curta molécula de RNA presente como um componente essencial de telomerase. A sequência telômero mostrada é a de humanos.



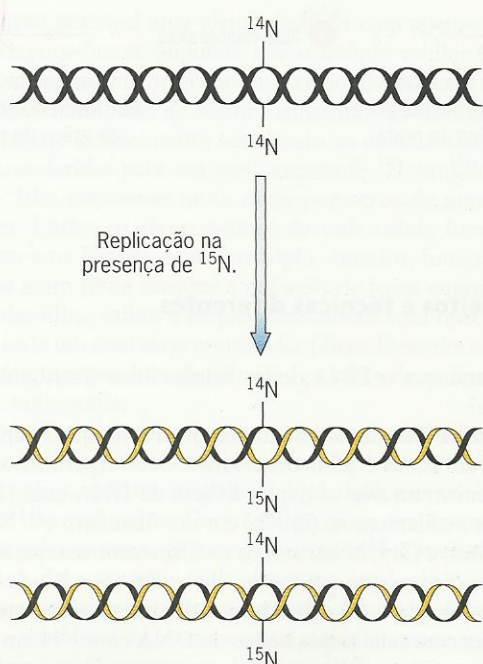
Fig. 10.34 ■ John Tacket, com 15 anos de idade, de Bay City, Michigan, fala sobre sua doença, a progeria, durante uma conferência em Washington em 16 de abril de 2003, para anunciar a descoberta do gene que causa esta condição genética rara e fatal, caracterizada pelo aparecimento de envelhecimento acelerado. À direita, o Dr. Francis Collins, diretor do National Human Genome Research Institute.

Exercícios Básicos

Ilustram a análise genética básica.

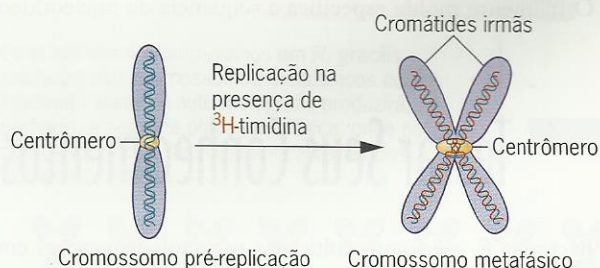
1. Bactérias *E. coli* que foram cultivadas em um meio normal contendo ^{14}N são transferidas para um meio contendo apenas o isótopo pesado de nitrogênio, ^{15}N , por uma geração de crescimento. Como o ^{14}N e o ^{15}N serão distribuídos no DNA destas bactérias após uma geração?

Resposta: Como o DNA se replica semiconservativamente, os filamentos parentais de DNA contendo ^{14}N serão conservados e usados como moldes para sintetizar novos filamentos complementares contendo ^{15}N . Assim, cada dupla hélice de DNA conterá um filamento leve e um filamento pesado, como mostrado no diagrama acompanhante.



2. Timidina radioativa (^3H) é adicionada ao meio de cultura no qual está sendo cultivada uma célula de camundongo. Esta célula não foi previamente exposta a nenhuma radioatividade. Se a célula estiver entrando na fase S no momento em que a ^3H -timidina é adicionada, que distribuição de radioatividade estará presente no DNA cromossômico na metáfase subsequente (a primeira metáfase após a adição de ^3H -timidina)?

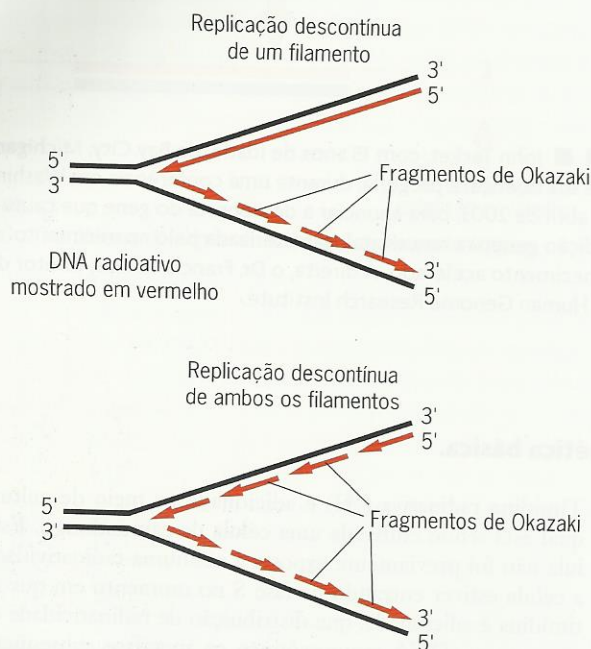
Resposta: Lembre que cada cromossomo pré-replicação contém uma única molécula grande de DNA que se estende de uma ponta do cromossomo através do centrômero até a outra ponta. Esta molécula de DNA irá replicar-se semiconservativamente como as moléculas de DNA em *E. coli* já discutidas. Entretanto, em metáfase, as duas duplas hélices da prole estarão presentes em cromátides irmãs ainda unidas no centrômero, como mostrado aqui:



3. Como a velocidade de centrifugação de gradiente de sacarose pode ser usada para distinguir entre (1) síntese descontínua de ambos os filamentos de DNA nascente em cada forquilha de replicação e (2) a replicação descontínua de um filamento e replicação contínua do outro filamento em cada forquilha?

Resposta: Usar ^3H -timidina para um DNA de pulso marcado em células que crescem em cultura e então extrair o DNA marcado, desnaturá-lo e determinar que proporção da marcação está pre-

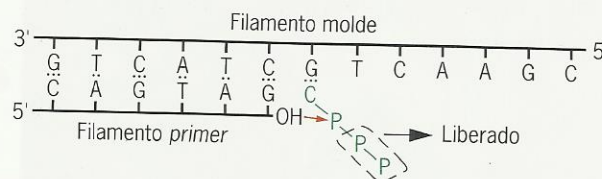
sente em fragmentos de Okazaki por velocidade de centrifugação de gradiente de sacarose (veja Enfoque Técnico: Técnicas de Centrifugação). Se ambos os filamentos estiverem se replicando descontinuamente, 100% do DNA de pulso marcado deve estar presente em fragmentos de Okazaki. Se apenas um filamento estiver se replicando descontinuamente, 50% da marcação deverá estar presente em fragmentos de Okazaki e os outros 50% em moléculas de DNA do tamanho do cromossomo. Estas duas possibilidades são diagramadas do seguinte modo:



4. DNA-polimerases só são capazes de sintetizar DNA na presença de um filamento molde e um filamento *primer*. Por quê? Quais são as funções destes dois filamentos?

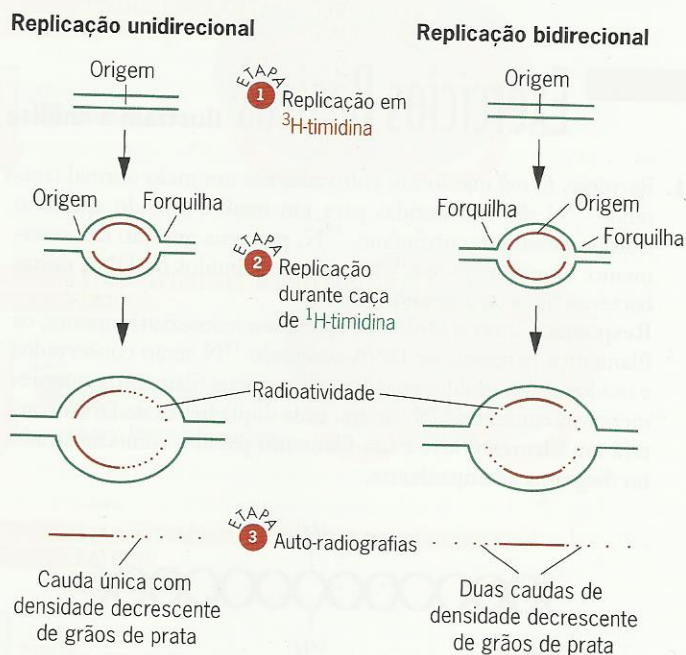
Resposta: DNA-polimerases só podem ampliar as cadeias de DNA com uma ponta 3'-OH livre porque o mecanismo de ampliação envolve um ataque nucleofílico pela 3'-OH no fósforo do interior do trifosfato de desoxirribonucleosídeo precursor com a eliminação de pirofosfato. O filamento com 3'-OH é o filamento *primer*. Ele é ampliado durante a síntese. O filamento molde especifica a sequência de nucleotídeos do

filamento que está sendo sintetizado; o novo filamento será complementar ao filamento molde. Estas funções são ilustradas do seguinte modo:



5. Como a auto-radiografia pode ser usada para distinguir entre replicação uni- e bidirecional do DNA?

Resposta: Se células são cultivadas em meio contendo ^3H -timidina por um curto período de tempo e são transferidas para um meio não radioativo para crescimento (um experimento de pulso-caça), a replicação uni- e bidirecional prevê padrões de marcação diferentes, e esses padrões podem ser diferenciados por auto-radiografia, como mostrado aqui:



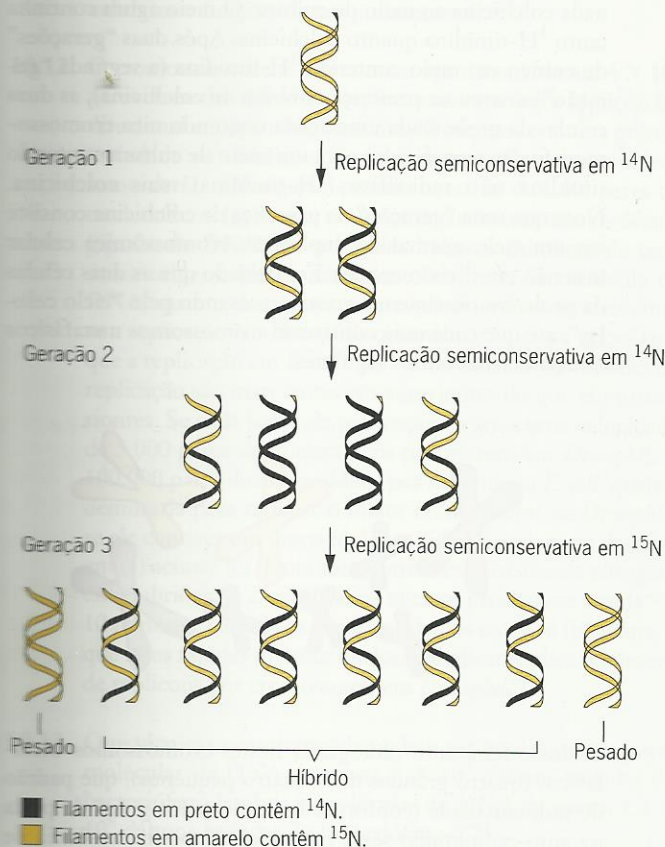
Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. Bactérias *E. coli* foram cultivadas por muitas gerações em um meio no qual o único nitrogênio disponível era o isótopo pesado ^{15}N . As células foram então coletadas por centrifugação, lavadas com um tampão e transferidas para um meio contendo ^{14}N (o isótopo de nitrogênio leve normal). Após duas gerações de crescimento em meio contendo ^{14}N , as bactérias foram transferidas de volta para um meio contendo ^{15}N para uma geração final de crescimento. Após esta geração final de crescimento na presença de ^{15}N , as células foram coletadas por centrifugação. O DNA foi então extraído destas células e analisado por centrifugação em gradiente de densidade de equilíbrio de CsCl. Como você

esperaria que o DNA destas células fosse distribuído no gradiente?

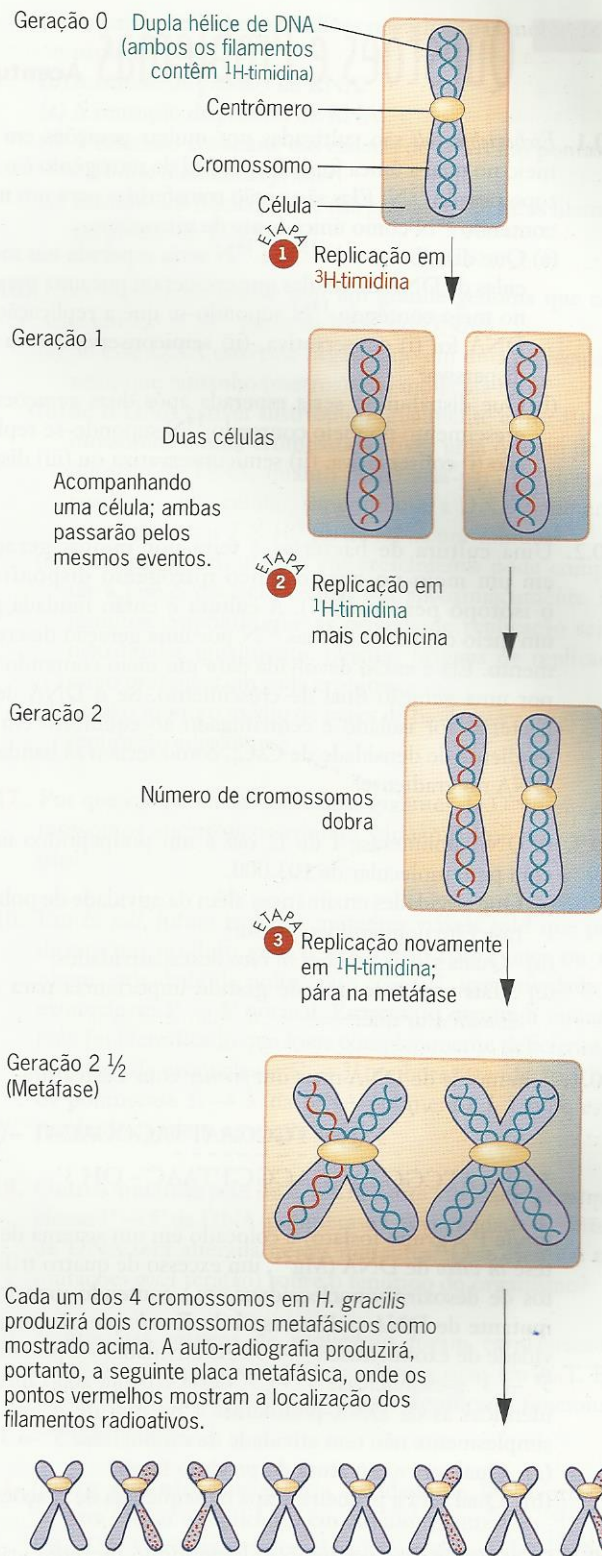
Resposta: Meselson e Stahl demonstraram que a replicação do DNA em *E. coli* é semiconservativa. Seus experimentos de controle mostram que as duplas hélices de DNA com (1) ^{14}N em ambos os filamentos, (2) ^{14}N em um filamento e ^{15}N no outro filamento e (3) ^{15}N em ambos os filamentos se separam em três bandas distintas no gradiente, chamadas (1) a banda leve, (2) a banda híbrida e (3) a banda pesada, respectivamente. Se você começa com uma dupla hélice de DNA com ^{15}N em ambos os filamentos e o replica semiconservativamente por duas gera-

ções na presença de ^{14}N , e então por uma geração na presença de ^{15}N , você acabará com oito moléculas de DNA, duas com ^{15}N em ambos os filamentos e seis com ^{14}N em um filamento e ^{15}N no outro filamento, como mostrado adiante. Portanto, 75% (6/8) do DNA aparecerá na banda híbrida e 25% (2/8) na banda pesada.



2. *Haplopappus gracilis* é uma planta diplóide com apenas dois pares de cromossomos. Suponha que o estágio celular G_1 desta planta, não previamente exposta a radioatividade de nenhum tipo, cresceu em meio de cultura contendo ^3H -timidina. Após uma geração de crescimento neste meio, as duas células da prole foram transferidas para um meio contendo ^1H -timidina e colchicina. Elas cresceram neste meio por cerca de uma e meia gerações. Então, os cromossomos de cada célula foram colocados em uma lâmina de microscópio, corados, fotografados e expostos a um filme sensível a radiação de baixa energia. Uma das células filhas exibiu uma placa metafásica com oito cromossomos, cada um com duas cromátides filhas. Desenhe essa placa metafásica mostrando a distribuição prevista de radioatividade na auto-radiografia.

Resposta: Cada cromossomo contém uma dupla hélice gigante que se replica semiconservativamente. No começo do experimento, ambos os filamentos dessa dupla hélice continham ^1H -timidina não radioativa. Como todos os quatro cromossomos passarão pelos mesmos processos de replicação, você só precisa seguir um deles e então multiplicar seu resultado por quatro. Como a colchicina não estava presente durante a primeira replicação, o número de cromossomos por célula permanece inalterado, mas os filamentos recém-sintetizados serão radioativos



(contêm ^3H -timidina). Como a segunda e terceira replicações ocorreram na presença de ^1H -timidina, filamentos sintetizados durante essas replicações serão não radioativos. Entretanto, a colchicina estava presente; logo, o número de cromossomos irá dobrar durante cada duplicação. Tais replicações são como mostrado na figura anexa, com filamentos radioativos de DNA mostrados em vermelho e filamentos não radioativos em azul.

Questões e Problemas Acentuam a compreensão e desenvolvem as habilidades analíticas.

10.1. *Escherichia coli* são cultivadas por muitas gerações em um meio no qual a única fonte disponível de nitrogênio é o isótopo pesado ^{15}N . Elas são então transferidas para um meio contendo ^{14}N como única fonte de nitrogênio.

(a) Que distribuição de ^{15}N e ^{14}N seria esperada nas moléculas de DNA de células que cresceram por uma geração no meio contendo ^{14}N supondo-se que a replicação do DNA foi (i) conservativa, (ii) semiconservativa ou (iii) dispersiva?

(b) Que distribuição seria esperada após duas gerações de crescimento no meio contendo ^{14}N supondo-se replicação (i) conservativa, (ii) semiconservativa ou (iii) dispersiva?

10.2. Uma cultura de bactérias é feita por muitas gerações em um meio no qual o único nitrogênio disponível é o isótopo pesado (^{15}N). A cultura é então mudada para um meio contendo apenas ^{14}N por uma geração de crescimento. Ela é então devolvida para um meio contendo ^{15}N por uma geração final de crescimento. Se o DNA destas bactérias for isolado e centrifugado ao equilíbrio em um gradiente de densidade de CsCl , como seriam as bandas de DNA no gradiente?

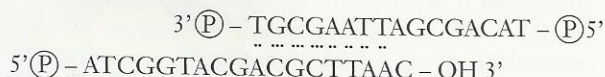
10.3. A DNA-polimerase I de *E. coli* é um polipeptídeo único com peso molecular de 103.000.

(a) Que atividades enzimáticas além da atividade de polimerase este polipeptídeo possui?

(b) Quais são as funções *in vivo* destas atividades?

(c) Tais atividades são de grande importância para uma *E. coli*? Por quê?

10.4. Um molde de DNA mais um *primer* com a estrutura



(onde P = grupo fosfato) é colocado em um sistema de síntese *in vitro* de DNA (Mg^{2+} , um excesso de quatro trifosfatos de desoxirribonucleosídeos etc.) contendo uma forma mutante de DNA-polimerase I de *E. coli* que não tem atividade de exonuclease $5' \rightarrow 3'$. As atividades de polimerase $5' \rightarrow 3'$ e exonuclease $3' \rightarrow 5'$ desta enzima aberrante são idênticas às da DNA-polimerase I normal de *E. coli*. Ela simplesmente não tem atividade de exonuclease $5' \rightarrow 3'$.

(a) Qual será a estrutura do produto final?

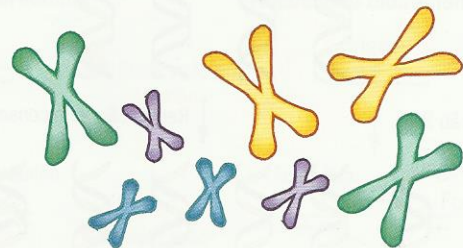
(b) Qual será a primeira etapa na sequência de reações?

10.5. Como os modos contínuo e descontínuo de replicação do DNA podem ser diferenciados experimentalmente?

10.6. Identifique as proteínas que estão envolvidas na replicação do DNA de *E. coli* e cite suas hipotéticas funções conhecidas.

10.7. A errante de Boston é uma planta imaginária com um número diplóide de cromossomos de 4. Suas células são facilmente cultivadas em culturas suspensas. Foi adicionada

^3H -timidina ao meio de cultura no qual o estágio celular G_1 desta planta estava crescendo. Após uma geração celular de crescimento em um meio contendo ^3H -timidina, foi adicionada colchicina ao meio de cultura. O meio agora continha tanto ^3H -timidina quanto colchicina. Após duas “gerações” de cultivo em meio contendo ^3H -timidina (a segunda “geração” ocorreu na presença também de colchicina), as duas células da prole (cada uma agora contendo oito cromossomos) foram transferidas para um meio de cultura contendo timidina não radioativa (^1H -timidina) mais colchicina. Note que uma “geração” na presença de colchicina consiste em um ciclo normal de duplicação cromossômica celular mas não em divisão celular. Foi deixado que as duas células da prole continuassem a crescer, passando pelo “ciclo celular”, até que cada uma contivesse cromossomos metafásicos que pareciam com os seguintes.



Se fosse feita auto-radiografia nestes cromossomos metafásicos (quatro grandes mais quatro pequenos), que padrão de radioatividade (conforme indicado pelos grãos de prata na auto-radiografia) seria esperado? (Suponha ausência de recombinação entre moléculas de DNA.)

10.8. Suponha que o experimento descrito no Problema 10.7 tenha sido feito novamente, exceto que desta vez substituindo a ^3H -timidina por timidina não radioativa ao mesmo tempo em que foi adicionada colchicina (após uma geração celular de crescimento em meio contendo ^3H -timidina). As células foram então mantidas em colchicina mais timidina não radioativa até ocorrer a metáfase mostrada no Problema 10.7. Qual o aspecto das auto-radiografias destes cromossomos?

10.9. Suponha que o DNA das células (mantidas em uma cultura) de uma espécie eucariótica fosse marcado por um curto período de tempo pela adição de ^3H -timidina ao meio. Em seguida, suponha que a marcação foi removida e as células foram ressuspensas em meio não radioativo. Após um curto período de crescimento em um meio não radioativo, o DNA foi extraído dessas células, diluído, suavemente colocado em filtros e auto-radiografado. Se fossem observadas auto-radiografias do tipo

— — — — —

o que isto indicaria sobre a natureza da replicação do DNA nestas células? Por quê?

- 10.10.** Arrume as seguintes enzimas na ordem de sua ação durante a replicação do DNA em *E. coli*: (1) DNA-polimerase I, (2) DNA-polimerase III, (3) DNA-primase, (4) DNA-girase e (5) DNA-helicase.
- 10.11.** Foram caracterizadas sete DNA-polimerases distintas – α , β , γ , δ , ϵ , ζ e η – em mamíferos. Quais as localizações intracelulares e funções destas polimerases?
- 10.12.** O cromossomo de *E. coli* contém aproximadamente 4×10^6 pares de nucleotídeos e se replica como um único réplicon bidirecional em aproximadamente 40 minutos em uma grande variedade de condições de crescimento. O maior cromossomo de *D. melanogaster* contém cerca de 6×10^7 pares de nucleotídeos. (a) Se este cromossomo contiver uma molécula gigante de DNA que se replica bidirecionalmente de uma única origem situada exatamente no meio da molécula de DNA, quanto demoraria para replicar todo o cromossomo se a replicação em *Drosophila* ocorresse na mesma velocidade que a replicação em *E. coli*? (b) Na verdade, as velocidades de replicação são mais lentas em eucariontes do que em procariontes. Se cada bolha de replicação cresce a uma velocidade de 5.000 pares de nucleotídeos por minuto em *Drosophila* e 100.000 pares de nucleotídeos por minuto em *E. coli*, quanto demoraria para replicar o maior cromossomo de *Drosophila* se ele contiver um único réplicon bidirecional como descrito em (a) acima? (c) Durante as primeiras divisões de clivagem em embriões de *Drosophila*, os núcleos dividem-se a cada 9 a 10 minutos. Com base em seus cálculos em (a) e (b) acima, o que estas rápidas divisões nucleares indicam sobre o número de réplicons por cromossomo em *Drosophila*?
- 10.13.** Que técnicas experimentais podem ser usadas para separar moléculas de DNA de massa 3×10^7 daltons isoladas do bacteriófago lambda e moléculas de DNA de massa $1,3 \times 10^8$ daltons isoladas do bacteriófago T2?
- 10.14.** O cromossomo do bacteriófago lambda tem vários segmentos ricos em A-T que se desnaturam quando expostos a pH 11,05 por 10 minutos. Após tal desnaturação parcial, a forma linear da molécula de DNA lambda tem a estrutura mostrada na Fig. 10.7a. Após sua injeção em uma *E. coli*, a molécula de DNA lambda é convertida em uma molécula circular covalentemente fechada por pontes de hidrogênio entre suas pontas complementares unifilamentares e submetida à ação da DNA ligase. Ela então se replica como uma estrutura em forma de θ . Todo o cromossomo lambda tem 17,5 μm de tamanho. Ele tem uma única origem de replicação situada a 14,3 μm da ponta esquerda da forma linear mostrada na Fig. 10.7a. Desenhe a estrutura que seria observada por microscopia eletrônica após tanto (1) replicação de um segmento com aproximadamente 6 μm de tamanho da molécula de DNA cromossômico lambda (*in vivo*) quanto (2) exposição desta molécula de DNA parcialmente replicada a pH 11,05 por 10 minutos (*in vitro*), (a) se a replicação tivesse ocorrido bidirecionalmente a partir da origem e (b) se a replicação tivesse ocorrido unidirecionalmente a partir da origem.
- 10.15.** Que atividade enzimática catalisa cada uma das seguintes etapas na replicação semiconservativa do DNA em procariontes?
- (a) A formação de super-hélices negativas nas moléculas da prole de DNA.
 (b) A síntese de *primers* de RNA.
 (c) A remoção de *primers* de RNA.
 (d) A extensão covalente das cadeias de DNA nas pontas 3'-OH dos filamentos *primer*.
 (e) Revisão dos nucleotídeos nas pontas 3'-OH dos filamentos *primer* de DNA?
- 10.16.** Uma espécie de árvore tem um grande genoma que consiste em $2,0 \times 10^{10}$ pares de bases do DNA.
 (a) Se esse DNA estivesse organizado em uma molécula linear, que tamanho (metros) esta molécula teria?
 (b) Se o DNA estiver distribuído igualmente entre 10 cromossomos e cada cromossomo tiver uma origem de replicação de DNA, quanto demoraria para completar a fase S do ciclo celular, supondo que a DNA-polimerase possa sintetizar 2×10^4 pb de DNA por minuto?
 (c) Uma célula ativamente em crescimento pode completar a fase S do ciclo celular em aproximadamente 300 minutos. Supondo que as origens de replicação sejam distribuídas igualmente, quantas origens de replicação estão presentes em cada cromossomo?
 (d) Qual o número médio de pares de bases entre origens de replicação adjacentes?
- 10.17.** Por que cada uma das moléculas gigantes de DNA em cromossomos eucarióticos contém várias origens de replicação?
- 10.18.** Em *E. coli*, foram isolados mutantes viáveis *polA* que produzem um produto gênico defeituoso com pouca ou nenhuma atividade de polimerase $5' \rightarrow 3'$, mas atividade de exonuclease $5' \rightarrow 3'$ normal. Entretanto, nenhum mutante *polA* foi identificado que fosse completamente deficiente de atividade de exonuclease $5' \rightarrow 3'$, embora retendo atividade de polimerase $5' \rightarrow 3'$ da DNA-polimerase I. Como estes resultados podem ser explicados?
- 10.19.** Outros mutantes *polA* de *E. coli* não têm atividade de exonuclease $3' \rightarrow 5'$ da DNA-polimerase I. A velocidade da síntese de DNA será alterada nestes mutantes? Que efeito(s) tais mutações *polA* terá(ão) sobre o fenótipo do organismo?
- 10.20.** Muitas das origens de replicação foram caracterizadas como possuidoras de seqüências cernes ricas em A-T. Estes cernes ricos em A-T têm algum significado funcional? Caso sim, qual?
- 10.21.** (a) Por que não é necessária atividade de DNA-primase para iniciar a replicação em círculo rolante?
 (b) A DNA-primase é necessária para a síntese descontínua do filamento *lagging*, que ocorre na cauda unifilamentar do círculo rolante. Por quê?
- 10.22.** A DNA-polimerase I é necessária para remover *primers* de RNA durante a replicação do cromossomo em *E. coli*. Entretanto, a DNA-polimerase III é a verdadeira replicase em *E. coli*. Por que a DNA-polimerase III não remove os *primers* de RNA?

- 10.23.** Em *E. coli*, são necessárias três proteínas diferentes para deslicoidizar a dupla hélice parental e manter desenrolados os filamentos em um molde distendido. O que são estas proteínas e quais suas respectivas funções?
- 10.24.** O quanto são similares as estruturas da DNA-polimerase I e da DNA-polimerase III em *E. coli*? Qual a estrutura da holoenzima DNA-polimerase III? Qual a função do produto do gene *dnaN* em *E. coli*?
- 10.25.** O produto do gene *dnaA* de *E. coli* é necessário para o início das síntese de DNA em *oriC*. Qual a sua função? Como sabemos que a proteína DnaA é essencial para a iniciação do processo?
- 10.26.** O que é um primossomo e quais são as suas funções? Que enzimas essenciais estão presentes no primossomo? Quais são os componentes principais do replissomo de *E. coli*? Como geneticistas podem determinar se estes componentes são necessários para replicação do DNA?
- 10.27.** O DNA cromossômico de eucariontes é acondicionado em nucleossomos durante a fase S do ciclo celular. Que obstáculos o tamanho e a complexidade tanto do replissomo quanto do nucleossomo apresentam durante a replicação semiconservativa do DNA eucariótico? Como estes obstáculos podem ser superados?
- 10.28.** Duas linhagens mutantes de *E. coli* têm cada uma uma mutação sensível a temperatura em um gene que codifica um produto necessário para duplicação cromossômica. Ambos os filamentos replicam seu DNA e se dividem normalmente a 25°C, mas são incapazes de replicar seu DNA ou de se dividir a 42°C. Quando células de uma linhagem são mudadas de crescimento a 25°C para 42°C, a síntese de DNA pára imediatamente. Quando células das outras linhagens são submetidas à mesma mudança de temperatura, a síntese de DNA continua, embora a uma velocidade menor, por cerca de meia hora. O que você pode concluir sobre as funções dos produtos destes dois genes?
- 10.29.** De que modos a replicação do DNA cromossômico em eucariontes difere da replicação do DNA em procariontes?
- 10.30.** (a) O cromossomo da bactéria *Salmonella typhimurium* contém cerca de 4×10^6 pares de nucleotídeos. Aproximadamente quantos fragmentos de Okazaki são produzidos durante uma replicação completa do cromossomo de *S. typhimurium*?
(b) O maior cromossomo de *D. melanogaster* contém aproximadamente 6×10^7 pares de nucleotídeos. Cerca de quantos fragmentos de Okazaki são produzidos durante a replicação deste cromossomo?
- 10.31.** Na levedura *S. cerevisiae*, células haplóides portadoras de uma mutação chamada *est1* (de telômeros cada vez mais curtos) perdem as seqüências distais teloméricas durante cada divisão celular. Preveja o efeito fenotípico final desta mutação na prole de tais células.
- 10.32.** A seqüência de uma dupla hélice de DNA é a seguinte:
- 5'-GATTCCCGGGGAAGCTTTGGGGGGCCCATCTTCGTACGTCTTTGCA-3'
3'-CTAAGGGGCCCTTTCGAACCCCGGGGTAGAAGCATGCAGAAACGT-5'
- Supondo que uma polimerase eucariótica possa replicar este pequeno pedaço de DNA, qual seria sua hipótese quanto ao que pareceriam os produtos da replicação?

Transcrição e Processamento do RNA

Tópicos

- O Controle Genético do Metabolismo: Uma Visão Geral
- Transferência da Informação Genética: O Dogma Central
- O Processo de Expressão Gênica
- Transcrição em Procariontes
- Transcrição e Processamento de RNA em Eucariontes
- Genes Interrompidos em Eucariontes: Éxons e Íntrons
- Remoção de Íntrons por Recomposição do RNA



Spliceossomo processando um transcrito gênico.

Estocagem e Transmissão da Informação com Códigos Simples

Vivemos na era do computador. Ele tem um impacto em praticamente todos os aspectos de nossas vidas, desde dirigir até o trabalho de observar naves espaciais pousando na lua. Estas mágicas eletrônicas podem estocar, recuperar e analisar dados com a velocidade da luz. O “cérebro” do computador é um pequeno *chip* de silício, o microprocessador, que contém um arranjo sofisticado e integrado de circuitos eletrônicos capazes de responder quase instantaneamente a impulsos codificados de energia elétrica. Efetuando coisas incríveis, o computador usa um código binário, uma linguagem baseada em 0 e 1. Assim, o alfabeto usado pelos computadores é como o do código Morse (pontos e traços) usado na telegrafia. Ambos consistem em apenas dois símbolos, em marcante contraste com as 26 letras do alfabeto inglês.

Obviamente, se o computador pode fazer suas mágicas com um alfabeto binário, grandes quantidades de informações podem ser estocadas e recuperadas sem usar códigos complexos ou alfabetos extensos. Neste e no capítulo seguinte, examinaremos (1) como a informação genética dos seres vivos é escrita em um alfabeto com apenas quatro letras, os quatro pares de bases no DNA, e (2) como essa informação genética é expressa durante o crescimento e o desenvolvimento de um organismo. Veremos que o RNA tem um papel central no processo da expressão gênica.

O CONTROLE GENÉTICO DO METABOLISMO: UMA VISÃO GERAL

Os genes codificam enzimas, e as enzimas catalisam os processos metabólicos em organismos vivos.

A capacidade dos organismos vivos de crescer e se reproduzir depende de um grande número de reações químicas. Os organismos devem sintetizar muitos tipos diferentes de moléculas das quais são compostos e degradar outras moléculas para obter energia utilizada no crescimento. As plantas têm a capacidade de usar energia da luz do sol para a síntese de macromoléculas, mas os animais devem obter energia dos alimentos que ingerem. Essa energia é obtida por quebra de moléculas grandes em moléculas menores e conversão da energia derivada deste processo em energia química estocada. Todas as reações que ocorrem em organismos vivos coletivamente são chamadas de *metabolismo*, e uma molécula orgânica que é sintetizada ou degradada é denominada *metabólito*. Nos quatro capítulos seguintes, examinaremos um quadro detalhado do controle genético do metabolismo e apresentaremos evidências que documentam as características mais importantes deste quadro. Começaremos com uma breve revisão do controle genético do metabolismo, que deve ajudar-nos a integrar as várias partes do quadro à medida que cada componente é adicionado.

Nos anos 1850, Louis Pasteur estudou a conversão de açúcar em álcool por “fermentos” em leveduras. Os fermentos de Pasteur foram depois chamados de *enzimas*, e, em 1897, Eduard Buchner extraiu enzimas de células de levedura, permitindo que tais “catalisadores” sejam estudados *in vitro*.

Um *catalisador* é uma substância que permite que uma reação ocorra sem ser modificada no processo. As enzimas catalisam reações metabólicas ligando-se fortemente a moléculas específicas, chamadas *substratos*, e as mantêm em justaposição, de modo que a energia necessária para a reação, a *energia de ativação*, seja reduzida. O resultado final é que ligações covalentes podem ser rearranjadas em temperaturas biológicas, as temperaturas do corpo de organismos vivos.

Embora lisados celulares fossem amplamente usados para estudar reações catalisadas por enzimas durante a primeira quarta parte do século vinte, a natureza química das enzimas permaneceu aberta a debates. Então, em 1926, James Sum-

ner extraiu a enzima urease de caroços de jaca, purificou-a na forma cristalina e demonstrou que ela consistia inteiramente em um tipo de macromolécula chamada *proteína*. Sabe-se que proteínas contêm constituintes chamados *aminoácidos* desde o início dos anos 1900, mas somente nos anos 1950, quando a estrutura da pequena proteína insulina foi determinada, é que a complexidade de cada proteína foi avaliada. Aminoácidos são pequenas moléculas orgânicas com uma gama de estruturas químicas, mas todos contêm um grupo amino ($-NH_2$) e um grupo carboxila ($-COOH$). Os 20 aminoácidos diferentes presentes em proteínas fornecem-lhes uma grande diversidade estrutural. Uma proteína pode conter uma única cadeia de aminoácidos, chamada *polipeptídeo*, ou duas ou mais cadeias (Cap. 12).

Milhares de enzimas já foram purificadas. A maioria é de proteínas com atividade catalítica, mas algumas são mais complexas, com componentes adicionais. As enzimas exibem especificidade marcante – cada enzima catalisa uma reação ou algumas reações muito similares. Elas variam de tamanho de cerca de 12.000 a mais de 1 milhão de daltons. A especificidade de cada enzima resulta de sua sequência única dos 20 aminoácidos. De fato, foi essa variabilidade essencialmente ilimitada de estruturas de proteína que levou muitos geneticistas a acreditar que as proteínas, e não os ácidos nucleicos, estocavam a informação genética.

A sequência de aminoácidos de cada polipeptídeo é controlada por um gene, e cada polipeptídeo tem um papel estrutural, regulatório ou enzimático na célula. Além disso, cada enzima catalisa uma reação metabólica específica. Entretanto, o quadro fica mais complexo neste ponto porque os processos metabólicos raramente envolvem apenas uma reação catalisada por enzima. Em vez disso, a síntese ou a degradação de um metabólito específico em geral ocorre por uma série de reações catalisadas por enzimas. O conjunto completo de reações catalisadas por enzimas necessárias para sintetizar ou degradar um determinado metabólito é chamado uma *via metabólica*. Por exemplo, considere a biossíntese da arginina, um dos 20 aminoácidos presentes na maioria das proteínas. Em *E. coli*, a síntese de arginina ocorre em oito etapas, cada uma catalisada por uma enzima específica (Fig. 11.1). As oito enzimas são compostas de nove produtos gênicos; uma enzima, a ornitina-carbamoyltransferase, contém dois produtos gênicos diferentes (ArgF e ArgI).

Em geral, cada etapa de uma via metabólica requer a atividade de uma enzima, e cada enzima contém um ou mais

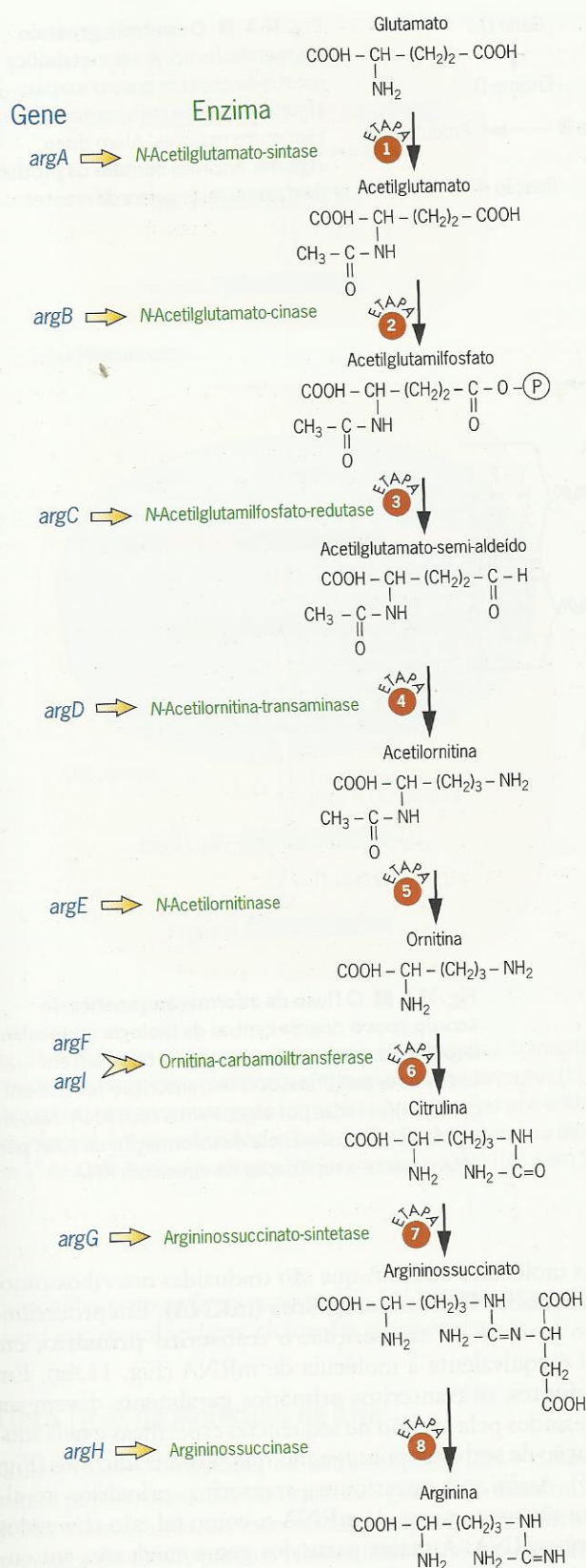


Fig. 11.1 ■ O controle genético da via biossintética de arginina em *E. coli*. Nove genes codificam oito enzimas que catalisam as oito etapas desta via metabólica.

polipeptídeos. Como cada polipeptídeo é especificado por um gene, cada uma das etapas em uma via metabólica precisará do produto de pelo menos um gene. Um quadro geral de controle das vias metabólicas é apresentado na Fig. 11.2. Embora a maioria dos genes codifiquem proteínas, os produtos finais de alguns genes são moléculas de RNA. Várias destas moléculas de RNA têm papéis essenciais na síntese de proteínas. Consideraremos suas funções em detalhes no Cap. 12. Como os genes controlam as estruturas dos RNA e de proteínas, devemos perguntar em seguida como as seqüências de pares de nucleotídeos nas moléculas de DNA especificam as seqüências de nucleotídeos no RNA e de aminoácidos nas moléculas de proteína.

PONTOS IMPORTANTES

- O metabolismo ocorre por seqüências de reações catalisadas por enzimas, com cada enzima especificada por um ou mais genes.

TRANSFERÊNCIA DA INFORMAÇÃO GENÉTICA: O DOGMA CENTRAL

O dogma central da biologia é que a informação estocada no DNA é transferida para moléculas de RNA durante a transcrição e para proteínas durante a tradução.

De acordo com o dogma central da biologia molecular, a informação genética geralmente flui (1) de DNA para DNA durante sua transmissão de geração a geração e (2) do DNA para proteína durante a expressão fenotípica em um organismo (Fig. 11.3). Durante a replicação de vírus com RNA, a informação também é transmitida de RNA para RNA. A transferência de informação genética do DNA para proteína envolve duas etapas: (1) **transcrição**, a transferência de informação genética do DNA para o RNA, e (2) **tradução**, a transferência de informação do RNA para proteína. Além disso, a informação genética flui do RNA para o DNA durante a conversão dos genomas dos vírus tumorais com RNA para suas formas pró-virais com DNA (Cap. 18). Assim, a transferência de informação genética do DNA para o RNA às vezes é reversível, enquanto a transferência da informação do RNA para proteína é sempre irreversível.

Transcrição e Tradução

Como já discutido, a expressão da informação genética ocorre em duas etapas: transcrição e tradução (Fig. 11.3). Durante a transcrição, um filamento de DNA de um gene é usado como um molde para sintetizar um filamento complementar de RNA, denominado **transcrito** gênico. Por exemplo, na Fig. 11.3, o filamento de DNA contendo a seqüência de nucleotídeos AAA é usado como um molde para produzir a

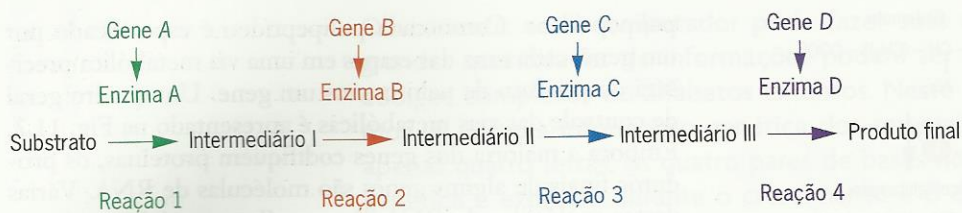


Fig. 11.2 ■ O controle genético do metabolismo. A via metabólica mostrada contém quatro etapas; algumas vias são mais curtas e outras são muito maiores. Além disso, algumas enzimas contêm os produtos de dois ou mais genes diferentes.

O dogma central

Fluxo de informação genética:

1. Perpetuação de informação genética de geração a geração

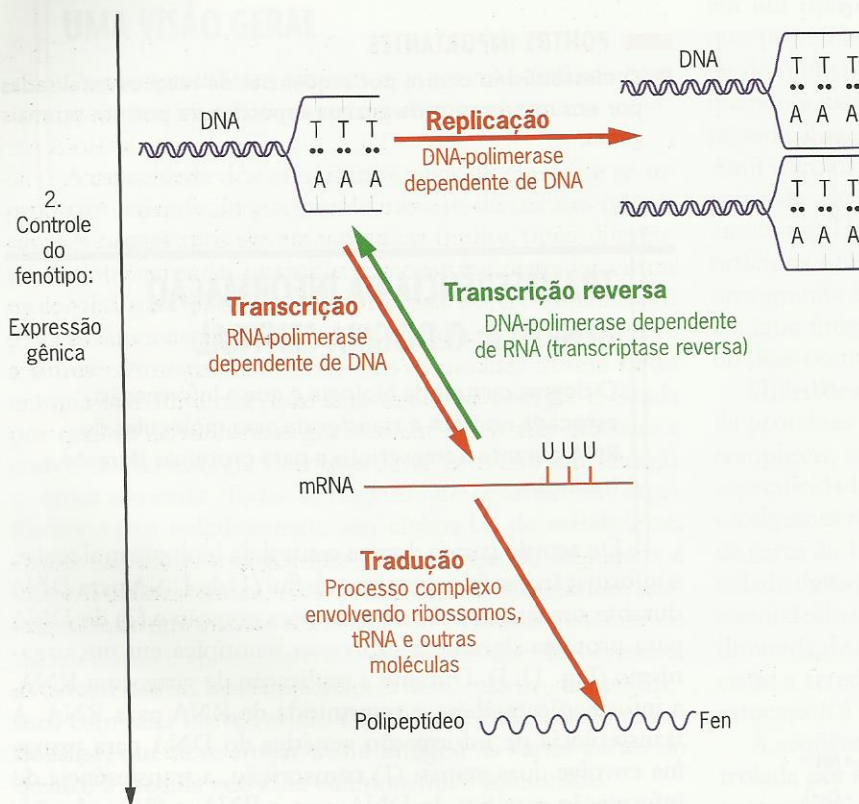


Fig. 11.3 ■ O fluxo da informação genética de acordo com o dogma central da biologia molecular. Replicação, transcrição e tradução ocorrem em todos os organismos; ocorre transcrição reversa em células infectadas por alguns vírus com RNA. Não é mostrada a transferência de informação de RNA para RNA durante a replicação de vírus com RNA.

seqüência complementar UUU no RNA transcrito. Durante a tradução, a seqüência de nucleotídeos no RNA transcrito é convertida na seqüência de aminoácidos do polipeptídeo produzido pelo gene. Essa conversão é governada pelo **código genético**, a especificação de aminoácidos por trincas de nucleotídeos chamadas **códons** no transcrito gênico. Por exemplo, a trinca UUU no RNA transcrito mostrado na Fig. 11.3 especifica o aminoácido fenilalanina (Fen) no polipeptídeo produzido pelo gene. A tradução ocorre em máquinas macromoleculares complexas chamadas **ribossomos**, que são compostas de três a cinco moléculas de RNA e 50 a 90 proteínas diferentes. Entretanto, o processo de tradução também requer a participação de muitas outras macromoléculas. Este capítulo enfoca transcrição; tradução é o assunto do Cap. 12.

As moléculas de RNA que são traduzidas nos ribossomos são chamadas **RNA mensageiros (mRNA)**. Em procariontes, o produto de transcrição, o **transcrito primário**, em geral é equivalente à molécula de mRNA (Fig. 11.4a). Em eucariontes, os transcritos primários geralmente devem ser processados pela excisão de seqüências específicas e pela modificação de ambas as pontas antes que sejam traduzidos (Fig. 11.4b). Assim, em eucariontes, transcritos primários geralmente são precursores de mRNA e, como tal, são chamados de **pré-mRNA**. A maior parte dos genes nucleares em eucariontes superiores e alguns eucariontes inferiores contêm seqüências não codificantes chamadas **íntrons** que separam as seqüências expressas ou **éxons** destes genes. Seqüências inteiras destes genes *divididos* são transcritas em pré-mRNA, e as seqüências íntron não codificantes são subsequentemente

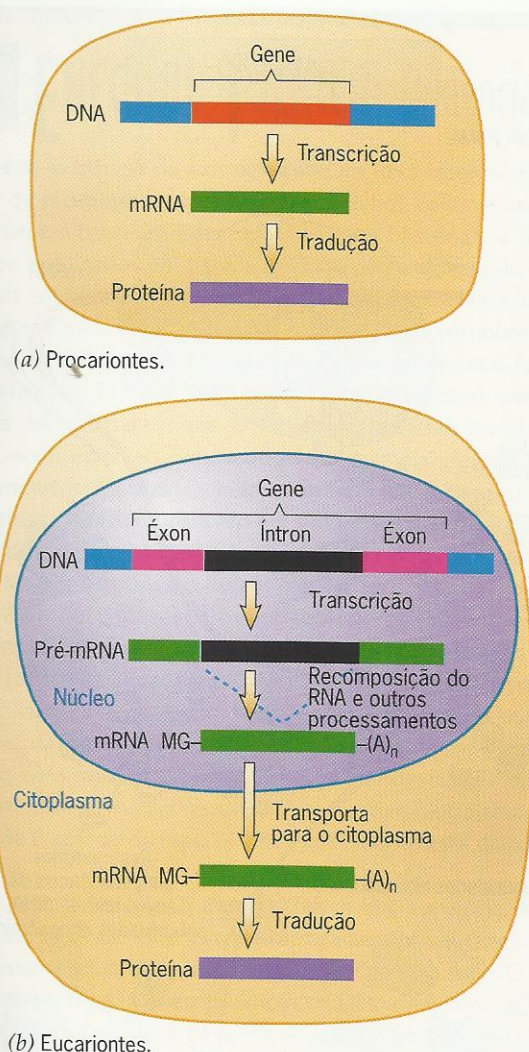


Fig. 11.4 ■ A síntese de proteínas envolve duas etapas: transcrição e tradução, tanto em procariontes (a) quanto em eucariontes (b). Além disso, em eucariontes, os transcritos primários ou pré-mRNA em geral devem ser processados pela remoção de íntrons e a adição em 5' de Cap. 7-metilguanossina (MG) e caudas poli(A) [(A)_n] em 3'.

removidas por *reações de recomposição (splicing)* feitas em estruturas macromoleculares chamadas *spliceossomos*.

Cinco Tipos de Moléculas de RNA

Cinco classes diferentes de moléculas de RNA desempenham papéis essenciais na expressão gênica. Já discutimos os RNA mensageiros, os intermediários que levam informação genética do DNA para os ribossomos onde são sintetizadas as proteínas. Os **RNA transportadores (tRNA)** são pequenas moléculas de RNA que funcionam como adaptadores entre aminoácidos e os códons no mRNA durante a tradução. Os **RNA ribossômicos (rRNA)** são componentes estruturais e catalíticos dos ribossomos, máquinas complexas

que traduzem as seqüências de nucleotídeos dos mRNA em seqüências de aminoácidos de polipeptídeos. Os **pequenos RNA nucleares (snRNA)** são componentes estruturais dos spliceossomos, estruturas nucleares que removem íntrons de genes nucleares. Os **micro-RNA (miRNA)** são RNA unifilamentares curtos, com 20 a 22 nucleotídeos, que são cortados de precursores pequenos em forma de grampo e bloqueiam a expressão de mRNA complementares ou parcialmente complementares causando sua degradação ou reprimindo sua tradução. Os papéis dos mRNA e snRNA são discutidos neste capítulo. As estruturas e funções dos tRNA e rRNA serão discutidas em detalhe no Cap. 12. Os mecanismos pelos quais os miRNA regulam a expressão gênica são discutidos no Cap. 21.

Todos os cinco tipos de RNA – mRNA, tRNA, rRNA, snRNA e miRNA – são produzidos por transcrição. Ao contrário dos mRNA, que especificam polipeptídeos, os produtos finais do tRNA, do rRNA, do snRNA e do miRNA são moléculas de RNA. As moléculas de RNA transportador, do RNA ribossômico, do snRNA e do miRNA não são traduzidas. A Fig. 11.5 mostra uma visão geral da síntese de proteínas em eucariontes, destacando a origem transcricional e funções dos cinco tipos de moléculas de RNA. O processo é similar em procariontes. Entretanto, em procariontes, o DNA não é separado dos ribossomos por um envoltório nuclear. Além disso, genes procarióticos raramente contêm seqüências não codificantes que são removidas durante o processamento do RNA transcrito.

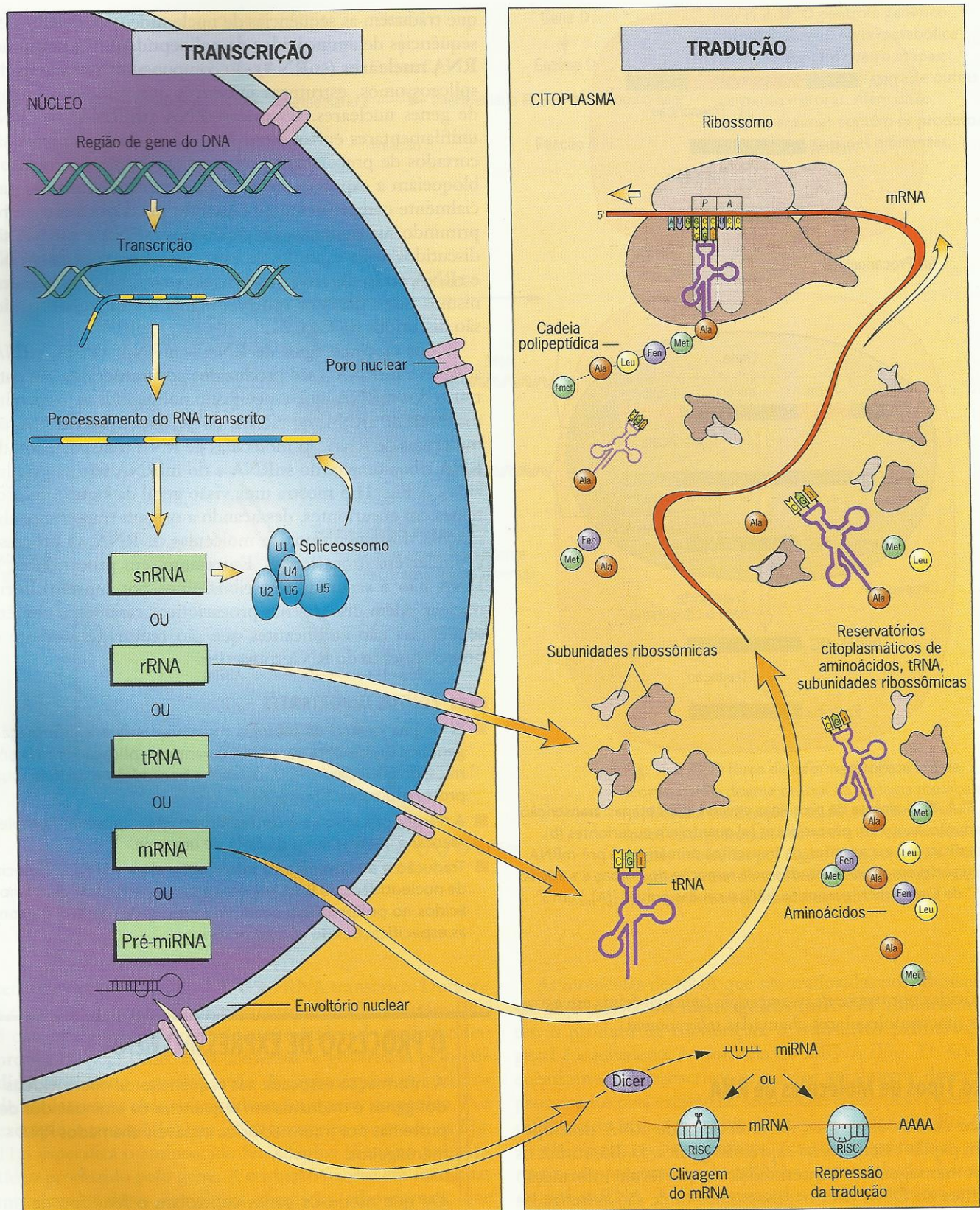
PONTOS IMPORTANTES

- O dogma central da biologia molecular é que a informação genética flui de DNA para DNA durante a replicação cromossômica, do DNA para o RNA durante a transcrição e do RNA para proteína durante a tradução.
- A transcrição envolve a síntese de um RNA transcrito complementar a um filamento de DNA de um gene.
- Tradução é a conversão de informação estocada na seqüência de nucleotídeos no RNA transcrito para a seqüência de aminoácidos no polipeptídeo produzido pelo gene, de acordo com as especificações do código genético.

O PROCESSO DE EXPRESSÃO GÊNICA

A informação estocada nas seqüências de nucleotídeos dos genes é traduzida em seqüências de aminoácidos de proteínas por intermediários instáveis chamados RNA mensageiros.

De que modo os genes controlam o fenótipo de um organismo? Como as seqüências de nucleotídeos dos genes dirigem o crescimento e o desenvolvimento de uma célula, um tecido, um órgão ou todo um ser vivo? Os geneticistas sabem que o fenótipo de um organismo é produzido pelos efeitos combinados de todos os seus genes que agem dentro



(a) Transcrição e processamento do RNA ocorrem no núcleo.

(b) Tradução ocorre no citoplasma.

Fig. 11.5 ■ Uma visão geral da síntese de proteínas, destacando a origem transcripcional de miRNA, snRNA, tRNA, rRNA e mRNA, a função de recomposição de snRNA, a regulação da expressão gênica por miRNA e os papéis traducionais de tRNA, rRNA, mRNA e ribossomos. Dicer é uma nuclease que processa o precursor de miRNA em miRNA, e RISC é o complexo silenciador induzido por RNA.



ENFOQUE TÉCNICO: Evidência de um RNA Mensageiro Instável

A primeira evidência da existência de um RNA intermediário na síntese de proteínas veio de estudos de Elliot Volkin e Lawrence Astrachan em bactérias infectadas por vírus bacterianos. Seus resultados, publicados em 1956, sugeriram que a síntese de proteínas virais em bactérias infectadas envolvia moléculas instáveis de RNA especificadas por DNA viral. Volkin e Astrachan observaram um surto de síntese de RNA após infecção de bactérias *E. coli* com o bacteriófago T2. Marcando o RNA com o isótopo radioativo ^{32}P , eles demonstraram que as moléculas de RNA recém-sintetizadas eram instáveis, renovando-se com meia-vida de apenas alguns minutos. Além disso, eles mostraram que a composição de nucleotídeos dos RNA instáveis era similar à composição do DNA de T2 e diferente da do DNA de *E. coli*. Seus resultados logo foram ampliados por estudos em outros laboratórios.

Em 1961, Sol Spiegelman e colaboradores relataram que os RNA instáveis sintetizados em células infectadas pelo fago T4 poderiam formar dúplexes RNA-DNA com DNA de T4 desnaturado, mas não com DNA de *E. coli* desnaturado. Eles fizeram pulsos de marcação das bactérias com ^3H -uridina em vários momentos após infecção com fago T4, isolaram o RNA total destas células e determinaram se as moléculas de RNA radioativo hibridizavam com o DNA de *E. coli* ou do fago T4. Seu experimento está diagramado na Fig. 1.

Seus resultados (Fig. 2) demonstraram que a maioria das moléculas de RNA de vida curta sintetizadas após infecção eram complementares a filamentos únicos do DNA do fago T4 e não complementares a filamentos únicos do DNA de *E. coli*, indicando que elas foram produzidas a partir de moldes do DNA do fago T4, e não de moldes do DNA de *E. coli*.

No mesmo ano em que Spiegelman e colaboradores publicaram seus resultados, Sydney Brenner, François Jacob e Matthew Meselson demonstraram que as proteínas do fago T4 eram sintetizadas em ribossomos de *E. coli*. Assim, as seqüências de aminoácidos das proteínas T4 não foram controladas por componentes dos ribossomos. Em vez disso, os ribossomos eram as bancadas de trabalho nas quais ocorria a síntese de proteínas, mas não deram as especificações para proteínas individuais. Estes resultados fortaleceram a idéia, primeiro formalmente proposta por François Jacob e Jacques Monod em 1961, de que moléculas instáveis de RNA levavam as especificações para seqüências de aminoácidos de produtos gênicos individuais dos genes aos ribossomos. Pesquisas subsequentes estabeleceram firmemente o papel desses RNA instáveis, hoje chamados de RNA mensageiros ou mRNA, na transferência de informação genética de genes para os sítios de síntese de proteínas no citoplasma.

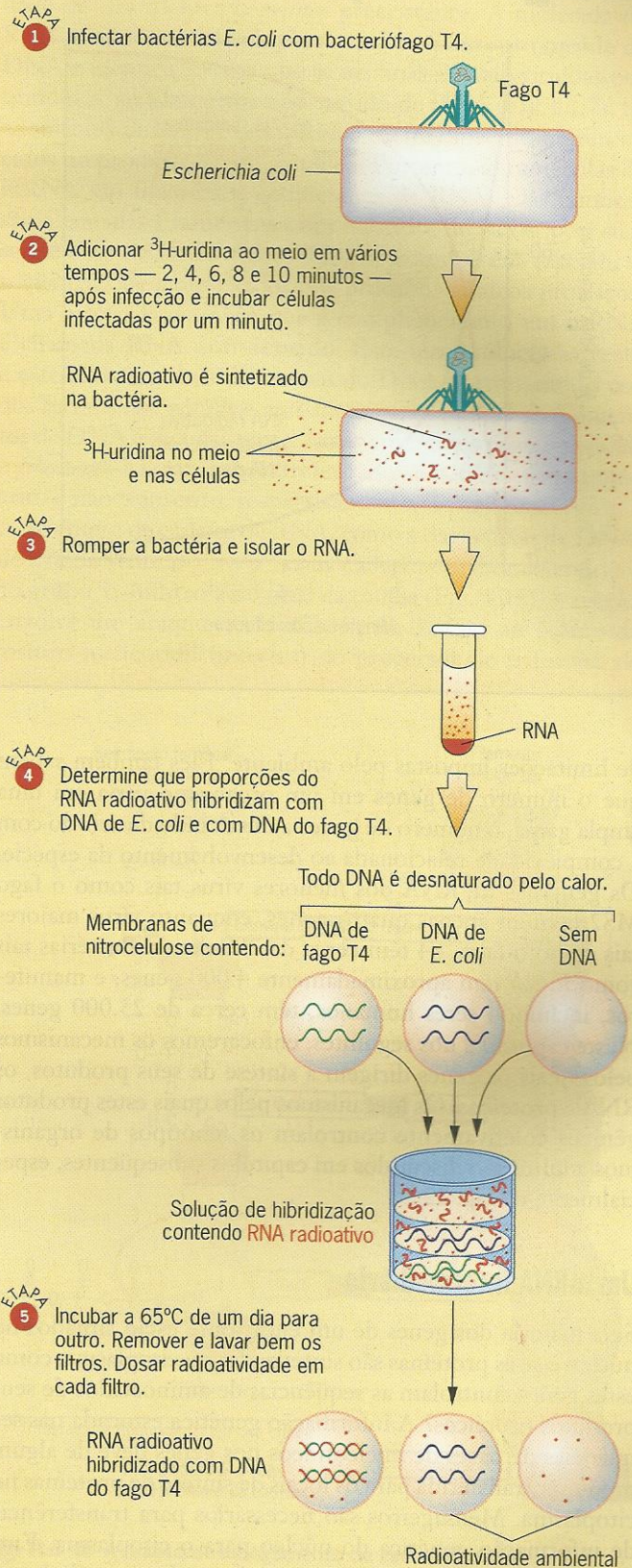


Fig. 1 ■ O experimento de Spiegelman.



ENFOQUE TÉCNICO: Evidência de um RNA Mensageiro Instável (continuação)

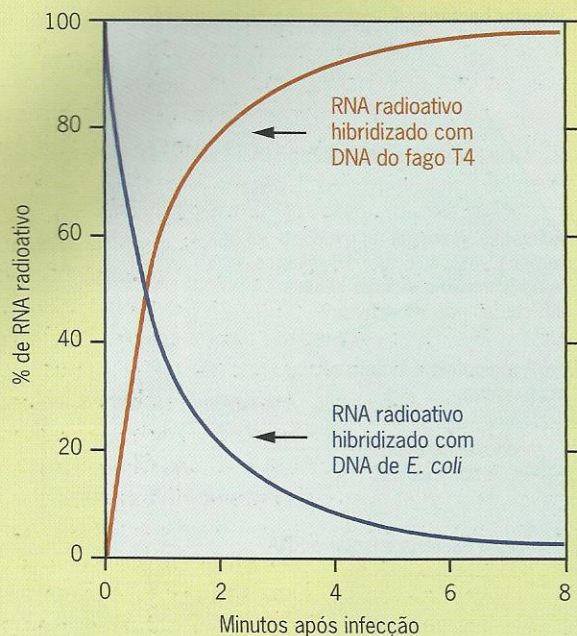


Fig. 2 ■ Mudança rápida da transcrição de genes de *E. coli* para genes do fago T4 em bactéria infectada pelo T4.

de limitações impostas pelo ambiente. Eles também sabem que o número de genes em um organismo varia em uma ampla gama, o número de genes aumentando de acordo com a complexidade relacionada ao desenvolvimento da espécie. Os genomas de RNA dos menores vírus tais como o fago MS2 contêm apenas quatro genes, enquanto vírus maiores tais como o fago T4 têm cerca de 200 genes. Bactérias tais como *E. coli* têm aproximadamente 4.000 genes, e mamíferos, incluindo seres humanos, têm cerca de 25.000 genes. Neste capítulo e nos seguintes, enfocaremos os mecanismos pelos quais os genes dirigem a síntese de seus produtos, os RNA e proteínas. Os mecanismos pelos quais estes produtos gênicos coletivamente controlam os fenótipos de organismos adultos são discutidos em capítulos subsequentes, especialmente o Cap. 22.

Um mRNA Intermediário

Se a maioria dos genes de um eucarionte estão situados no núcleo e se as proteínas são sintetizadas no citoplasma, como estes genes controlam as seqüências de aminoácidos de seus produtos protéicos? A informação genética estocada nas seqüências de pares de nucleotídeos nos genes deve de algum modo ser transferida para os locais de síntese de proteínas no citoplasma. Mensageiros são necessários para transferência de informação genética do núcleo para o citoplasma. Embora a necessidade de tais mensageiros seja mais óbvia em

eucariontes, a primeira evidência de sua existência veio de estudos de procariontes (veja Enfoque Técnico: Evidência de um RNA Mensageiro Instável).

A síntese de RNA mensageiros ou precursores de mRNA nos núcleos de eucariontes e seu subsequente transporte para o citoplasma pode ser documentada por experimentos de marcação de pulso, experimentos de pulso-caça e autoradiografia. Se uma célula que cresce em cultura é exposta a um precursor de RNA radioativo tal como ^3H -uridina ou ^3H -citidina por alguns minutos e a localização intracelular da radioatividade incorporada é determinada por autoradiografia, o RNA marcado está presente quase que exclusivamente no núcleo (Fig. 11.6a). Entretanto, se a curta exposição ao precursor radioativo for seguida de um período de crescimento em meio não radioativo, a maioria da radioatividade incorporada estará presente no citoplasma (Fig. 11.6b). Assim, os RNA intermediários instáveis são sintetizados no núcleo e são transportados para o citoplasma, onde dirigem a síntese de proteínas. Neste capítulo, enfocaremos a síntese, o processamento e o transporte destas moléculas de RNA mensageiro.

Características Gerais da Síntese de RNA

A síntese de RNA ocorre por um mecanismo que é similar ao da síntese de DNA (Cap. 10) exceto que (1) os precursores são **trifosfatos de ribonucleosídeos** e não trifosfatos de

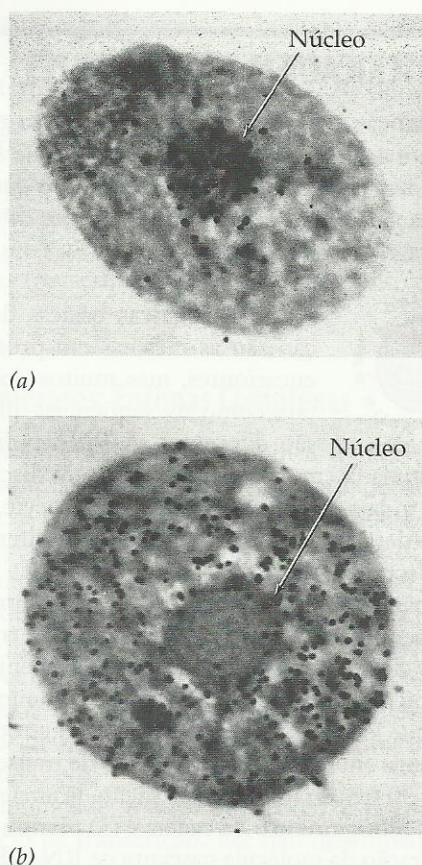


Fig. 11.6 ■ Auto-radiografias demonstrando a síntese de RNA no núcleo e seu subsequente transporte para o citoplasma. Cada auto-radiografia é superposta sobre uma foto de um corte fino da célula. Os pontos pretos representam grânulos de prata na emulsão auto-radiográfica que reagiram com elétrons emitidos pelo decaimento dos átomos de ^3H . (a) Uma célula de *Tetrahymena* marcada com ^3H -citidina por 15 minutos. (b) Uma célula de *Tetrahymena* que foi cultivada em meio não radioativo por 88 minutos após exposição a ^3H -citidina por 12 minutos.

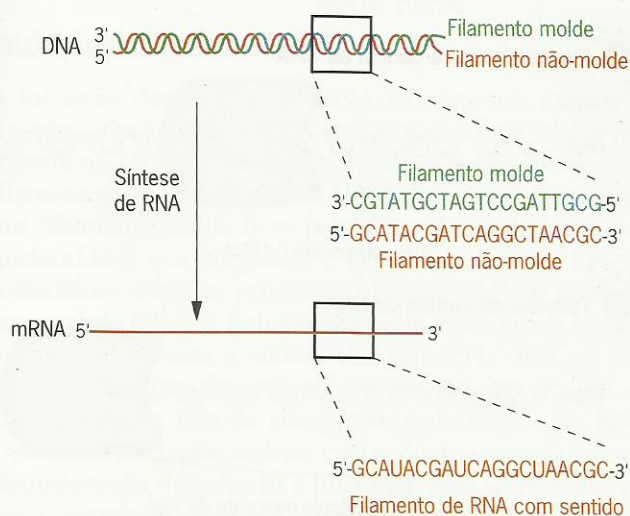


Fig. 11.7 ■ A síntese de RNA usa um filamento de DNA de um gene como molde.

desoxirribonucleosídeos, (2) apenas um filamento de DNA é usado como molde para a síntese de uma cadeia de RNA complementar em determinada região e (3) as cadeias de RNA podem ser iniciadas *de novo*, sem nenhuma necessidade de um filamento *primer* preexistente. A molécula de RNA produzida será complementar ao **filamento molde** de DNA e idêntica, exceto que as uridinas são substituídas por timidinas, ao **filamento não-molde** de DNA (Fig. 11.7). Se a molécula de RNA é um mRNA, ela irá especificar aminoácidos no produto gênico protéico. Portanto, as moléculas de mRNA são filamentos codificantes de RNA. Elas também são chamadas **filamentos com sentido**, de RNA, pois suas seqüências de nucleotídeos “fazem sentido” em especificar seqüências de aminoácidos nos produtos gênicos protéicos. Uma molécula de RNA que é complementar a um mRNA é chamada **RNA anti-sentido**. Esta terminologia às vezes é estendida aos dois filamentos de DNA. Entretanto, o uso dos termos *com sentido* e *anti-sentido* para indicar os filamentos de DNA é inconsistente. Assim, usaremos *filamento molde* e *filamento não-molde* para referir-nos aos filamentos transcrito e não transcrito, respectivamente, de um gene.

A síntese de cadeias de RNA, como a das cadeias de DNA, ocorre no sentido $5' \rightarrow 3'$, com a adição de ribonucleotídeos ao grupo 3'-hidroxila no final da cadeia (Fig. 11.8). A reação envolve um ataque nucleofílico pela 3'-OH ao átomo de fósforo nucleotídil (interior) do precursor de trifosfato de

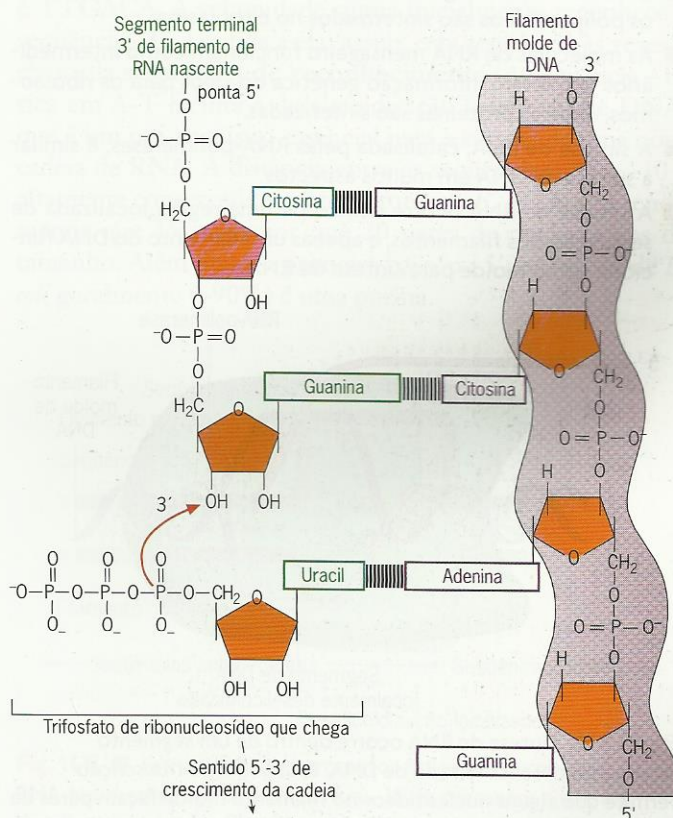


Fig. 11.8 ■ A reação de alongamento de cadeia do RNA catalisada pela RNA-polimerase.

Ao discutir a transcrição, os biólogos geralmente usam os termos *upstream* (antecedente) e *downstream* (posterior) para se referir às regiões que estão para a ponta 5' e ponta 3', respectivamente, do transcrito de um ponto na molécula de mRNA. Estes termos são baseados no fato de que a síntese de RNA sempre ocorre no sentido de 5' para 3'. As regiões antecedente e posterior dos genes são as seqüências de DNA que especificam os segmentos correspondentes 5' e 3' de seus transcritos em relação a um ponto específico de referência.

RNA-polimerases: Enzimas Complexas

As RNA-polimerases que catalisam a transcrição são proteínas complexas, multiméricas. A RNA-polimerase de *E. coli* tem peso molecular de cerca de 480.000 e consiste em cinco polipeptídeos. Dois deles são idênticos; portanto, a enzima contém quatro polipeptídeos distintos. A molécula completa de RNA-polimerase, a **holoenzima**, tem a composição $\alpha_2\beta\beta'\sigma$. As subunidades α estão envolvidas na montagem do **cerne tetramérico** ($\alpha_2\beta\beta'$) da RNA-polimerase. A subunidade β contém o sítio de ligação de trifosfato de ribonucleosídeo e a subunidade β' abriga a região de ligação ao molde de DNA.

Uma subunidade, o **fator sigma** (σ), está envolvido apenas no início da transcrição. Ele não desempenha papel no alongamento da cadeia. Após ter ocorrido início da cadeia de RNA, o fator σ é liberado, e o alongamento da cadeia (veja Fig. 11.8) é catalisado pela enzima cerne ($\alpha_2\beta\beta'$). A função de sigma é reconhecer e ligar a RNA-polimerase ao início da transcrição ou sítios promotores no DNA. A enzima cerne (sem σ) irá catalisar a síntese de RNA a partir de moldes de DNA *in vitro*, mas, ao fazer isto, irá iniciar cadeias de RNA em sítios aleatórios em ambos os filamentos de DNA. Em contraste, a holoenzima (σ presente) inicia cadeias de RNA *in vitro* apenas em sítios usados *in vivo*.

Iniciação das Cadeias de RNA

A iniciação das cadeias de RNA envolve três etapas: (1) ligação da holoenzima RNA-polimerase a uma região promotora no DNA; (2) a deselcoidização localizada dos dois filamentos de DNA pela RNA-polimerase, fornecendo um filamento molde livre para fazer pares de bases com nucleotídeos que cheguem; e (3) a formação de ligações fosfodiéster entre os primeiros ribonucleotídeos na cadeia nascente de RNA. A holoenzima permanece ligada à região promotora durante a síntese das primeiras oito ou nove ligações. Então, o fator sigma é liberado, e o cerne da enzima começa a fase de alongamento da síntese de RNA. Durante a iniciação, cadeias curtas de dois a nove ribonucleotídeos são sintetizadas e liberadas. Essa síntese abortiva cessa quando tiverem sido sintetizadas cadeias com 10 ou mais ribonucleotídeos e a RNA-polimerase tiver começado a se afastar do promotor.

Por convenção, os pares de nucleotídeos dentro de unidades de transcrição e adjacentes a elas são numerados em relação ao início da transcrição (designada como +1) – o par de nucleotídeos correspondente ao primeiro nucleotídeo (5') do RNA transcrito. Os pares de bases que precedem o sítio de iniciação recebem prefixos menos (-); os seguintes (em relação ao sentido da transcrição) ao sítio de início recebem o sinal mais (+). As seqüências de nucleotídeos que precedem o sítio de iniciação são chamadas **seqüências antecedentes**; as que vêm depois do sítio de iniciação são chamadas **seqüências posteriores**.

Como já foi mencionado, a subunidade sigma da RNA-polimerase medeia sua ligação a promotores no DNA. Centenas de promotores de *E. coli* foram seqüenciados e têm surpreendentemente pouco em comum. Duas seqüências curtas dentro desses promotores são suficientemente conservadas para serem reconhecidas, mas mesmo estas raramente são idênticas em dois promotores diferentes. Os pontos médios das duas seqüências conservadas ocorrem a cerca de 10 e 35 pares de nucleotídeos, respectivamente, antes do sítio de início da transcrição (Fig. 11.11). Assim, elas são chamadas **seqüência -10** e **seqüência -35**, respectivamente. Embora tais seqüências variem um pouco de gene para gene, alguns nucleotídeos são altamente conservados. As seqüências de nucleotídeos que estão presentes em tais elementos genéticos conservados são mais freqüentemente chamadas de **seqüências de consenso**. A seqüência de consenso -10 no filamento não-molde é TATAAT; a seqüência de consenso -35 é TTGACA. A subunidade sigma inicialmente reconhece a seqüência -35 e se liga a ela; assim, esta seqüência às vezes é chamada **seqüência de reconhecimento**. A seqüência -10 rica em A-T facilita a deselcoidização localizada de DNA, que é um pré-requisito essencial para a síntese de uma nova cadeia de RNA. A distância entre as seqüências -35 e -10 é altamente conservada em promotores de *E. coli*, nunca sendo menor que 15 ou maior que 20 pares de nucleotídeos de tamanho. Além disso, a primeira base, ou 5', nos RNA de *E. coli* geralmente (>90%) é uma purina.

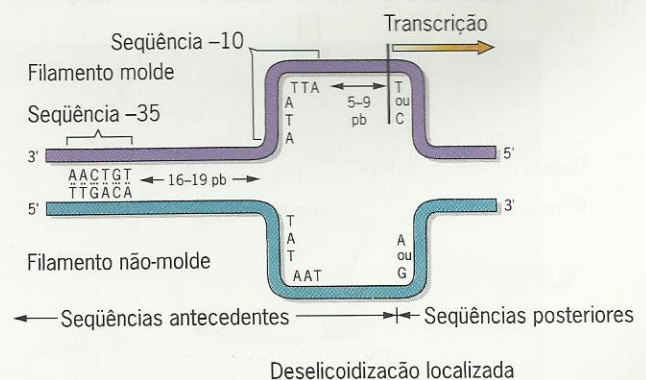
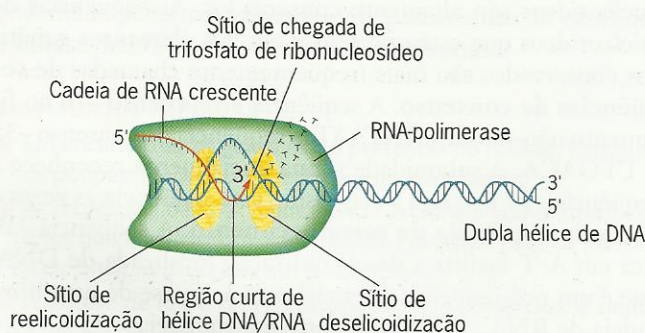


Fig. 11.11 ■ Estrutura de um promotor típico em *E. coli*. A RNA-polimerase liga-se à seqüência -35 do promotor e inicia a deselcoidização dos filamentos de DNA na seqüência -10 rica em A-T. A transcrição começa dentro da bolha de transcrição em um sítio de cinco a nove pares de bases além da seqüência -10.

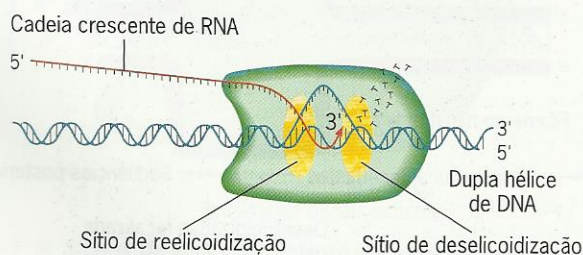
Alongamento das Cadeias de RNA

O alongamento das cadeias de RNA é catalisado pelo cerne da enzima RNA-polimerase, após a liberação da subunidade σ . A ampliação covalente das cadeias de RNA (veja Fig. 11.8) ocorre dentro da bolha de transcrição, um segmento localmente desenrolado do DNA. A molécula de RNA-polimerase contém atividades tanto de desenrolamento quanto de reelcoidização do DNA. A RNA-polimerase desenrola continuamente a dupla hélice de DNA à frente do sítio de polimerização e reenrola os filamentos complementares de DNA atrás do sítio de polimerização à medida que se movimenta ao longo da dupla hélice (Fig. 11.12). Em *E. coli*, o tamanho médio de uma bolha de transcrição é de 18 pares de nucleotídeos, e cerca de 40 ribonucleotídeos são incorporados na cadeia crescente de RNA por segundo. A cadeia nascente de RNA é deslocada do filamento molde de DNA à medida que a RNA-polimerase se move ao longo da molécula de DNA. A região de pareamento transitório de bases entre a cadeia crescente e o filamento molde de DNA é muito curta, talvez apenas três pares de bases de tamanho. A estabilidade do complexo de transcrição é mantida primaria-



(a) RNA-polimerase é ligada ao DNA e covalentemente amplia a cadeia de RNA.

Movimento da RNA-polimerase



(b) RNA-polimerase moveu-se para além de sua posição em (a), ampliando processivamente a cadeia de RNA nascente.

Fig. 11.12 ■ Alongamento de uma cadeia de RNA catalisada pela RNA-polimerase em *E. coli*.

mente pela ligação do DNA e da cadeia crescente de RNA à RNA-polimerase, e não pelo pareamento de bases entre o filamento molde de DNA e o RNA nascente.

Término das Cadeias de RNA

O término das cadeias de RNA ocorre quando a RNA-polimerase encontra um **signal de término**. Quando isto ocorre, o complexo de transcrição se dissocia, liberando a molécula nascente de RNA. Existem dois tipos de finalizadores de transcrição em *E. coli*. Um tipo resulta em término apenas na presença de uma proteína chamada *rô* (ρ). Portanto, tais seqüências finalizadoras são chamadas *finalizadores dependentes de rô*. O outro tipo resulta no término da transcrição sem o envolvimento de *rô*. Tais seqüências são chamadas *finalizadores independentes de rô*.

Os finalizadores independentes de *rô* contêm uma região rica em G-C seguida de seis ou mais pares de bases A-T, com os A presentes no filamento molde (Fig. 11.13, em cima). A seqüência de nucleotídeos da região rica em G-C é tal que regiões unifilamentares do RNA podem fazer pares de bases e formar estruturas em grampo (Fig. 11.13, embaixo). As estruturas em grampo do RNA formam-se imediatamente após a síntese das regiões participantes da cadeia de RNA e retardam o movimento das moléculas de RNA-polimerase ao longo do DNA, causando pausas na ampliação da cadeia. Como o pareamento de bases A-U é fraco, precisando de menos energia para separar as bases que qualquer outro par de bases padrão, a série de U após a região do grampo provavelmente facilita a liberação das cadeias de RNA recém-sintetizadas a partir do molde de DNA quando a estrutura em grampo faz com que a RNA-polimerase pare neste sítio.

O mecanismo pelo qual ocorre o término da transcrição dependente de *rô* ainda é incerto. As seqüências de término dependentes de *rô* têm de 50 a 90 pares de bases de comprimento e especificam RNA transcritos que são ricos em C e muito desprovidas de G. Além disso, sinais de término diferentes dependentes de *rô* têm pouco em comum. A proteína *rô* liga-se à cadeia crescente de RNA e se move de 5' para 3' ao longo do RNA, parecendo seguir a molécula de RNA-polimerase que catalisa a síntese da cadeia. Quando a RNA-polimerase diminui ou pára na seqüência de término dependente de *rô*, *rô* junta-se à polimerase e tira a cadeia nascente de RNA da bolha de transcrição.

Transcrição, Tradução e Degradação Concomitantes do mRNA

Em procariontes, a tradução e a degradação de uma molécula de mRNA em geral começam antes que sua síntese (transcrição) esteja completa. Como as moléculas de mRNA são sintetizadas, traduzidas e degradadas no sentido 5' para 3', todos os três processos podem ocorrer simultaneamente na mesma molécula de RNA. Em procariontes, a maquinaria de síntese de polipeptídeos não é separada por um envoltório-

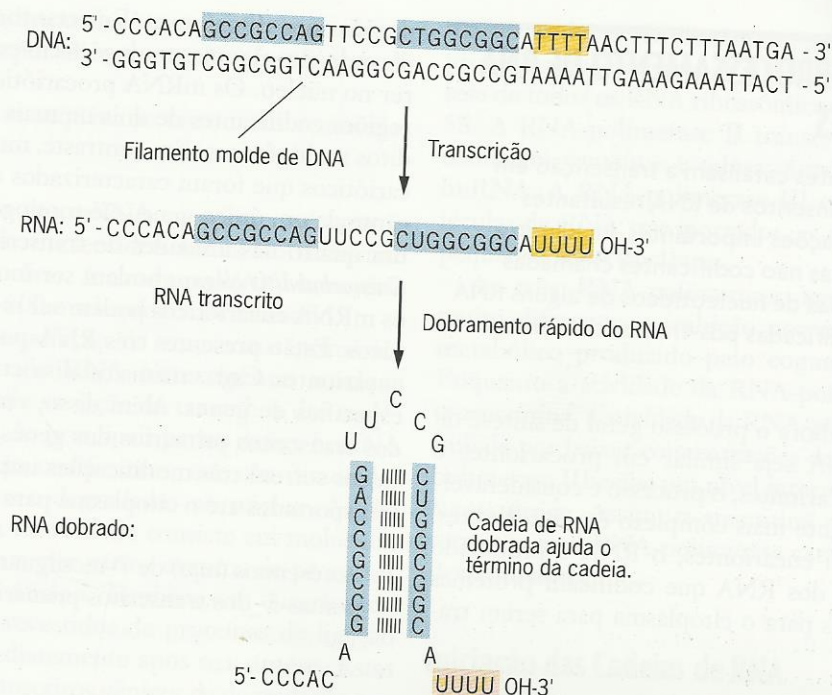


Fig. 11.13 ■ Estrutura de um finalizador de transcrição independente de ρ . As sequências finalizadoras independentes de ρ contêm uma região rica em G-C seguida de pelo menos seis pares de bases A-T. A transcrição de tal sequência finalizadora produz uma cadeia de RNA com segmentos ricos em G-C que fazem pares de bases uns com os outros imediatamente após a síntese para formar uma estrutura em grampo. Esta estrutura retarda o movimento da RNA-polimerase ao longo da molécula de DNA, o que resulta no término da transcrição na via A-T adjacente e na liberação da cadeia nascente de RNA.

rio nuclear do local de síntese do mRNA. Portanto, uma vez que tenha sido sintetizada a ponta 5' de um mRNA, ele pode ser imediatamente usado como molde para a síntese de polipeptídeos. De fato, a transcrição e a tradução em geral são muito acopladas em procariotes. Oscar Miller, Barbara Hamkalo e colaboradores desenvolveram técnicas que lhes permitem visualizar este acoplamento entre transcrição e tradução em bactérias por microscopia eletrônica. Uma de suas fotos mostrando a transcrição acoplada de um gene e a tradução de seu mRNA produzido em *E. coli* é mostrada na Fig. 11.14.

PONTOS IMPORTANTES

- A síntese de RNA ocorre em três estágios: (1) iniciação, (2) alongamento e (3) término.
- RNA-polimerases, as enzimas que catalisam a transcrição, são proteínas multiméricas complexas.
- A ampliação covalente das cadeias de RNA ocorre dentro de segmentos de DNA desenrolados localmente.
- O alongamento da cadeia pára quando a RNA-polimerase encontra um sinal de término de transcrição.
- Transcrição, tradução e degradação de moléculas de mRNA em geral ocorrem simultaneamente em procariotes.

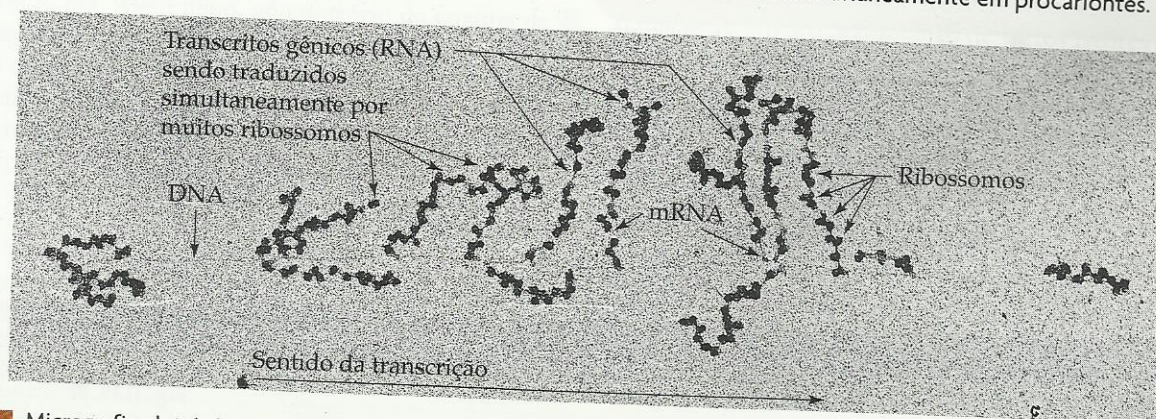


Fig. 11.14 ■ Micrografia eletrônica preparada por Oscar Miller e Barbara Hamkalo mostrando a transcrição e a tradução acoplada de um gene em *E. coli*. São visíveis o DNA, os mRNA e os ribossomos que traduzem moléculas individuais de mRNA. As cadeias polipeptídicas nascentes que estão sendo sintetizadas nos ribossomos não são visíveis à medida que se dobram em sua configuração tridimensional durante a síntese.

TRANSCRIÇÃO E PROCESSAMENTO DE RNA EM EUKARIOTES

Três enzimas diferentes catalisam a transcrição em eucariontes, e os transcritos de RNA resultantes sofrem três modificações importantes, incluindo a excisão de seqüências não codificantes chamadas íntrons. As seqüências de nucleotídeos de alguns RNA transcritos são modificadas pós-transcricionalmente por edição do RNA.



Embora o processo geral de síntese de RNA seja similar em procariontes e eucariontes, o processo é consideravelmente mais complexo em eucariontes. Em eucariontes, o RNA é produzido

no núcleo, e a maioria dos RNA que codificam proteínas devem ser transportados para o citoplasma para serem tra-

duzidos nos ribossomos. Entretanto, existem evidências indicando que algumas traduções em eucariontes podem ocorrer no núcleo. Os mRNA procarióticos em geral contêm as regiões codificantes de dois ou mais genes. Tais mRNA são ditos multigênicos. Em contraste, muitos dos transcritos eucarióticos que foram caracterizados contêm a região codificante de um único gene (são monogênicos). Entretanto, até um quarto das unidades de transcrição no pequeno verme *Caenorhabditis elegans* podem ser multigênicas. Claramente, os mRNA eucarióticos podem ser monogênicos ou multigênicos. Estão presentes três RNA-polimerases diferentes em eucariontes. Cada enzima catalisa a transcrição de uma classe específica de genes. Além disso, em eucariontes, a maioria dos transcritos primários dos genes que codificam polipeptídeos sofrem três modificações importantes antes de serem transportados até o citoplasma para tradução (Fig. 11.15).

1. Acréscimos (*caps*) de 7-metilguanossina são adicionados às pontas 5' dos transcritos primários.

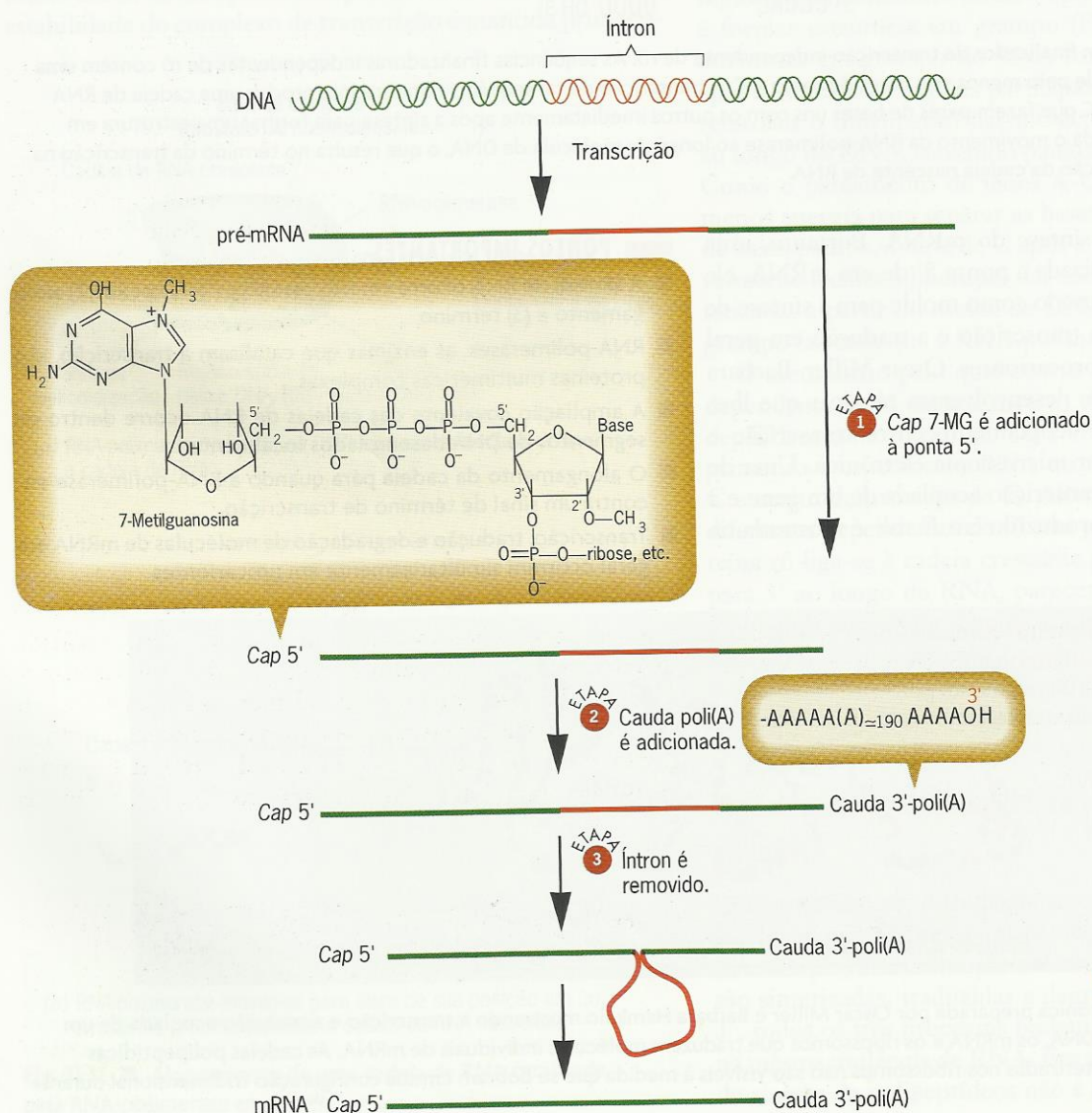


Fig. 11.15 ■ Em eucariontes, a maioria dos transcritos gênicos sofre três tipos diferentes de processamento pós-transcricional.

2. Caudas poli(A) são adicionadas às pontas 3' dos transcritos, que são geradas por clivagem e não por término da ampliação da cadeia.
3. Quando presentes, as seqüências íntron são removidas dos transcritos.

O **cap de 5'** na maioria dos mRNA eucarióticos é uma 7-metilguanosina unida ao nucleosídeo inicial do transcrito por uma ligação 5'-5' fosfato. A **cauda poli(A)** é uma série de adenosinas com 20 a 200 nucleotídeos de tamanho.

Em eucariontes, a população de transcritos primários em um núcleo é chamada **RNA heterogêneo nuclear (hnRNA)** devido à grande variação de tamanhos das moléculas de RNA presentes. As principais partes desses hnRNA são seqüências íntron não codificantes, que são removidas dos transcritos primários e degradadas no núcleo. Assim, grande parte do hnRNA na verdade consiste em moléculas de pré-mRNA que passam por vários eventos de processamento antes de deixar o núcleo. Também, em eucariontes, os RNA transcritos são revestidos de proteínas de ligação ao RNA durante ou imediatamente após sua síntese. Estas proteínas protegem os transcritos gênicos da degradação por ribonucleases, enzimas que degradam moléculas de RNA, durante processamento e transporte para o citoplasma. A vida média de um transcrito de um gene em eucariontes é de cerca de cinco horas, em contraste com a vida média de menos de cinco minutos em *E. coli*. Esta estabilidade acentuada de transcritos gênicos em eucariontes é dada, pelo menos em parte, por sua presença em complexos com proteínas de ligação ao RNA.

Três RNA-polimerases/Três Conjuntos de Genes

Enquanto uma única RNA-polimerase catalisa toda a transcrição em *E. coli*, os eucariontes, que variam em complexidade de uma só célula em leveduras até humanos, contêm três RNA-polimerases diferentes. Todas as três enzimas eucarióticas, chamadas **RNA-polimerases I, II e III**, são mais complexas, com 10 ou mais subunidades, que a RNA-polimerase de *E. coli*. Além disso, ao contrário da enzima de *E. coli*, todas as três RNA-polimerases eucarióticas precisam da ajuda de outras proteínas chamadas **fatores de transcrição** para iniciar a síntese de cadeias de RNA.

As principais características das três RNA-polimerases eucarióticas estão resumidas no Quadro 11.1. A RNA-polimerase I está situada no nucléolo, uma região distinta do

núcleo onde os rRNA são sintetizados e combinados com proteínas ribossômicas. A RNA-polimerase I catalisa a síntese de todos os RNA ribossômicos exceto o pequeno rRNA 5S. A RNA-polimerase II transcreve genes nucleares que codificam proteínas e talvez outros genes que especificam hnRNA. A RNA-polimerase III catalisa a síntese das moléculas de RNA transportador, as moléculas de rRNA 5S e pequenos RNA nucleares.

As três RNA-polimerases apresentam sensibilidades muito diferentes a inibição por α -amanitina, um veneno metabólico produzido pelo cogumelo *Amanita phalloides*. Enquanto a atividade da RNA-polimerase I é insensível a α -amanitina, a atividade da RNA-polimerase II é totalmente inibida por baixas concentrações de α -amanitina, e a RNA-polimerase III exibe um nível intermediário de sensibilidade a esta droga. Assim, α -amanitina pode ser usada para determinar qual RNA-polimerase catalisa a transcrição de um determinado gene.

Iniciação das Cadeias de RNA

Ao contrário de suas contrapartes procarióticas, as RNA-polimerases eucarióticas não podem iniciar a transcrição por si mesmas. Todas as três RNA-polimerases eucarióticas precisam da ajuda de fatores de transcrição protéicos para começar a síntese de uma cadeia de RNA. Estes fatores de transcrição devem ligar-se a uma região promotora no DNA e formar um complexo de iniciação apropriado antes que a RNA-polimerase se ligue e inicie a transcrição. São usados promotores e fatores de transcrição diferentes pelas RNA-polimerases I, II e III. Nesta seção, enfocaremos a iniciação da síntese de pré-mRNA pela RNA-polimerase II, que transcreve a maioria dos genes eucarióticos.

Em todos os casos, a iniciação da transcrição envolve a formação de um segmento localmente deselicoidizado de DNA, fornecendo um filamento de DNA livre para funcionar como um molde para a síntese de um filamento complementar de RNA (veja Fig. 11.9). A formação do segmento localmente deselicoidizado de DNA requerido para iniciar a transcrição envolve a interação de vários fatores de transcrição com seqüências específicas no promotor para a unidade de transcrição. Os promotores reconhecidos pela RNA-polimerase II consistem em curtos elementos conservados, ou módulos, situados antecedentes ao ponto inicial de transcrição. Os componentes do promotor do gene de timidina-cinase de camundongo são mostrados na Fig. 11.16. Outros

QUADRO 11.1

Características das Três RNA-polimerases de Eucariontes

Enzima	Localização	Produtos	Sensibilidade a α -Amanitina
RNA-polimerase I	Nucléolo	RNA ribossômicos, excluindo o 5S rRNA	Sem sensibilidade
RNA-polimerase II	Núcleo	Pré-mRNA nucleares	Sensibilidade completa
RNA-polimerase III	Núcleo	tRNA, 5S rRNA e outros RNA pequenos nucleares	Sensibilidade intermediária

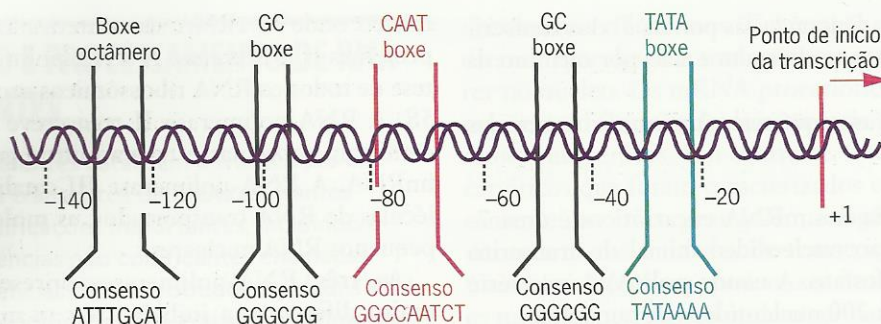


Fig. 11.16 ■ Estrutura de um promotor reconhecido pela RNA-polimerase II. Os TATA e CAAT boxes estão situados mais ou menos nas mesmas posições nos promotores da maioria dos genes nucleares que codificam proteínas. Os GC boxes e octâmeros podem estar presentes ou ausentes. Quando presentes, eles ocorrem em muitos locais diferentes, seja isoladamente ou em várias cópias. As seqüências mostradas aqui são as seqüências de consenso para cada um dos elementos promotores. Os elementos promotores conservados são mostrados em seus locais no gene de timidina-cinase de camundongo.

promotores que são reconhecidos pela RNA-polimerase II contêm parte desses componentes mas não todos eles. O elemento conservado mais próximo do sítio de início da transcrição (posição +1) é chamado **TATA boxe**. Ele tem a seqüência de consenso TATAAAA (lendo de 5' para 3' no filamento não-molde) e é centrado na posição -30. O TATA boxe tem um papel importante em posicionar o ponto de início da transcrição. O segundo elemento conservado é chamado **CAAT boxe**. Ele geralmente ocorre perto da posição -80 e tem a seqüência de consenso GGCCAATCT. Dois outros elementos conservados, o **GC boxe**, consenso GGGCGG, e o **boxe octâmero**, consenso ATTTGCAT, geralmente estão presentes nos promotores da RNA-polimerase II. Eles influenciam a eficiência de um promotor em iniciar a transcrição.

O início da transcrição pela RNA-polimerase II requer a assistência de vários **fatores básicos de transcrição**. Outros fatores de transcrição e seqüências regulatórias chamados de *acentuadores* e *silenciadores* modulam a eficiência da iniciação (Cap. 21). Os fatores básicos de transcrição devem interagir com promotores na seqüência correta para iniciar efetivamente a transcrição (Fig. 11.17). Cada fator basal de transcrição é indicado como **TFIIX** (fator de transcrição para a polimerase II), em que X é a letra que identifica o fator individual).

As posições nas quais cada um desses fatores de transcrição se liga ao DNA do promotor foram estabelecidas em parte determinando-se que segmentos de DNA estão protegidos da degradação por nucleases quando um fator específico foi adicionado ao complexo de iniciação. As seqüências de DNA que estão protegidas por fatores de transcrição ou outras proteínas de ligação ao DNA podem ser identificadas por experimentos de *footprinting* de DNase. Pequenas moléculas homogêneas de DNA que contêm o sítio potencial de ligação são isoladas e marcadas em uma ponta com um isótopo radioativo ou corante fluorescente. Metade das moléculas são incubadas com a proteína de ligação ao DNA de interesse e metade é mantida livre de proteínas. Ambas as

amostras de DNA são então rapidamente tratadas – apenas o suficiente para que a maioria das moléculas de DNA sejam clivadas uma vez – com DNase I pancreática, uma enzima que faz cortes internos aleatórios no DNA. Os produtos resultantes de clivagem são então separados com base no tamanho por eletroforese em gel (veja Cap. 9). A amostra livre de proteína conterá fragmentos de DNA de todos os tamanhos possíveis produzidos por clivagem entre todos os pares de nucleotídeos adjacentes na molécula. Os produtos de clivagem produzidos a partir do DNA ao qual a proteína estava ligada exibirão uma “pegada” (*footprint*) onde não existir ponta de fragmento de DNA. A seqüência à qual a proteína é ligada será protegida de clivagem pela DNase I. Comparando as bandas de DNA radioativo produzidas nas duas amostras, com e sem proteína de ligação ao DNA, um pesquisador pode definir com precisão a seqüência de nucleotídeos à qual a proteína se liga.

TFIID é o primeiro fator basal de transcrição a interagir com o promotor. Ele contém uma proteína de ligação a TATA (TBP) e várias pequenas proteínas associadas a TBP (Fig. 11.17). Em seguida, TFIIA junta-se ao complexo, seguido por TFIIB. TFIIF primeiro associa-se à RNA-polimerase II, e então TFIIF e RNA-polimerase II juntam-se ao complexo de iniciação da transcrição. TFIIF contém duas subunidades, uma das quais tem atividade de deselicoidear o DNA. Portanto, TFIIF provavelmente catalisa o deselicoidear localizado da dupla hélice de DNA necessário para iniciar a transcrição. TFIIE então se junta ao complexo de iniciação, ligando-se ao DNA posterior (*downstream*) do ponto de início da transcrição. Dois outros fatores, TFIIH e TFIIF, juntam-se ao complexo após TFIIE, mas suas localizações no complexo são desconhecidas.

As RNA-polimerases I e III iniciam a transcrição por processos que são similares, mas um pouco mais simples, que o usado pela polimerase II. Entretanto, os promotores dos genes transcritos pelas polimerases I e III são bem diferentes dos usados pela polimerase II, muito embora às vezes contenham alguns dos mesmos elementos de ação *cis*. Os

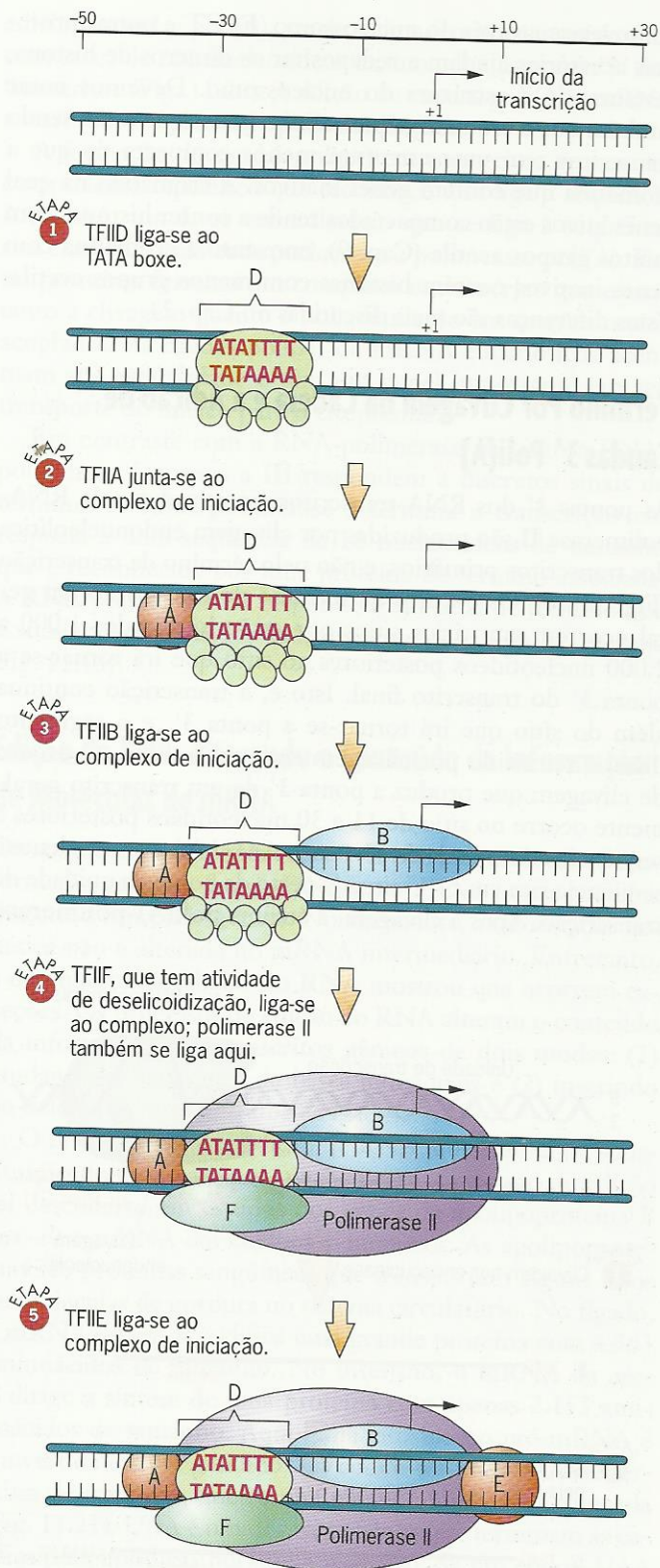


Fig. 11.17 ■ O início da transcrição pela RNA-polimerase II requer a formação de um complexo basal de início de transcrição na região promotora. A montagem deste complexo começa quando TFIID, que contém a proteína de ligação a TATA (TBP), liga-se ao TATA box. Os outros fatores de transcrição e a RNA-polimerase II juntam-se ao complexo na sequência mostrada.

promotores da RNA-polimerase I são bipartidos, com uma sequência cerne que se estende de cerca de -45 a +20 e um elemento antecedente de controle que vai de -180 a cerca de -105. As duas regiões têm sequências similares, e ambas são ricas de G-C. A sequência cerne é suficiente para a iniciação. Entretanto, a eficiência da iniciação é fortemente aumentada pela presença do elemento de controle antecedente.

Curiosamente, os promotores da maioria dos genes transcritos pela RNA-polimerase III estão situados dentro das unidades de transcrição, posteriores aos pontos de início da transcrição, e não antecedentes como nas unidades transcritas pelas RNA-polimerases I e II. Os promotores de outros genes transcritos pela polimerase III estão situados antecedentes ao ponto de início da transcrição, como para as polimerases I e II. De fato, os promotores da polimerase III podem ser divididos em três classes, duas das quais têm promotores situados dentro da unidade de transcrição.

Alongamento da Cadeia de RNA e a Adição de 5' Metilguanosina

Quando as RNA-polimerases eucarióticas são liberadas de seus complexos de iniciação, elas catalisam o alongamento da cadeia de RNA pelo mesmo mecanismo utilizado pelas RNA-polimerases de procariontes (veja Figs. 11.8 e 11.9).

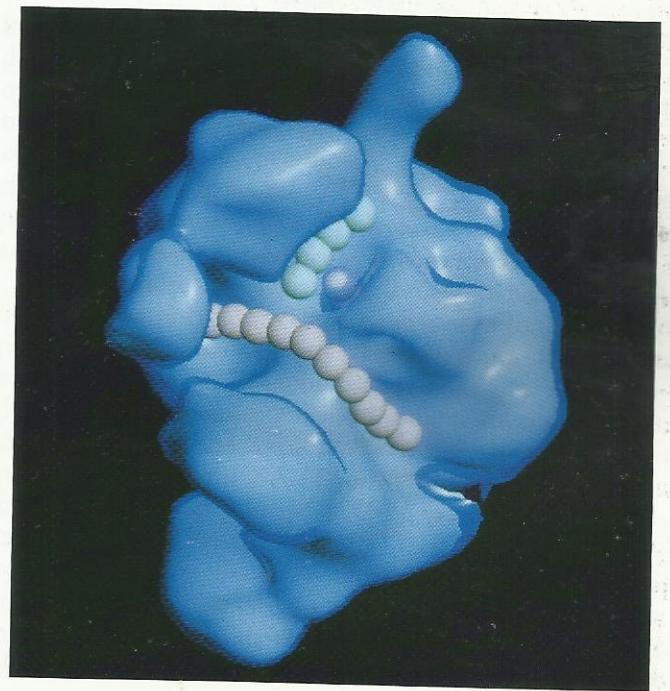


Fig. 11.18 ■ Modelo da estrutura da RNA-polimerase II na levedura *S. cerevisiae*. Os sulcos na superfície da enzima parecem ser sítios de ligação para o DNA e a cadeia crescente de RNA. Um sulco com cerca de 2,5 nm de largura e 1 nm de profundidade é o sítio hipotético de ligação para o DNA (a cadeia com 10 contos de tamanho); outro sulco com cerca de 1,5 nm de largura e 2 nm de profundidade liga-se à cadeia crescente de RNA (a cadeia com 4 contos de tamanho).

A análise estrutural da RNA-polimerase II de *S. cerevisiae* revela a presença de sulcos de superfície tidos como sendo sítios de ligação de DNA e RNA (Fig. 11.18).

No início do processo de alongamento, as pontas 5' dos pré-mRNA eucarióticos são modificadas pela adição de "revestimentos" (*caps*) de 7-metilguanosina (7-MG). Estes "revestimentos" 7-MG são adicionados quando as cadeias crescentes de RNA têm apenas cerca de 30 nucleotídeos de tamanho (Fig. 11.19). O 7-MG contém uma ligação incomum 5'-5'-trifosfato (veja Fig. 11.15) e dois ou mais grupos metila. Esses revestimentos de 5' são adicionados co-transcritionalmente pela via biossintética mostrada na Fig. 11.19. Os *caps* 7-MG são reconhecidos por fatores protéicos envolvidos na iniciação da transcrição (Cap. 12) e também ajudam a proteger as cadeias crescentes de RNA de degradação por nucleases.

Lembre que genes eucarióticos estão presentes na cromatina organizada em nucleossomos (Cap. 9). Como a RNA-polimerase transcreve DNA acondicionado em nucleossomos? Os nucleossomos têm que ser desmontados antes que o DNA dentro dele possa ser transcrito? Surpreendentemente, a RNA-polimerase II é capaz de mover-se através dos nucleossomos com a ajuda de um complexo protéico chamado FACT (facilitador de transcrição da cromatina), que remove dímeros de histonas H2A/H2B dos nucleossomos deixando "hexassomos" de histona. Após a polimerase

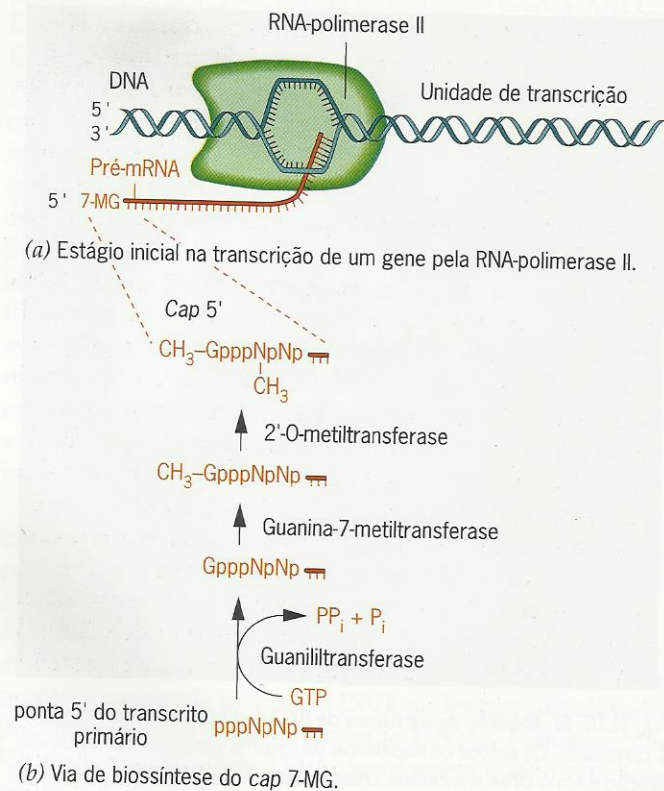


Fig. 11.19 ■ Os revestimentos de 7-metilguanosina (7-MG) são adicionados às pontas 5' dos pré-mRNA logo após começar o processo de alongamento.

II mover-se através do nucleossomo, FACT e outras proteínas acessórias ajudam a redepositar os dímeros de histona, restaurando a estrutura do nucleossomo. Devemos notar também que a cromatina que contém genes que estão sendo transcritos tem uma estrutura menos compacta do que a cromatina que contém genes inativos. A cromatina na qual genes ativos estão compactados tende a conter histonas com muitos grupos acetila (Cap. 9), enquanto a cromatina com genes inativos contém histonas com menos grupos acetila. Estas diferenças são mais discutidas no Cap. 21.

Término Por Clivagem da Cadeia e a Adição de Caudas 3' Poli(A)

As pontas 3' dos RNA transcritos sintetizados pela RNA-polimerase II são produzidas por clivagem endonucleolítica dos transcritos primários, e não pelo término da transcrição (Fig. 11.20). Os eventos de término de transcrição em geral ocorrem em vários sítios que estão localizados 1.000 a 2.000 nucleotídeos posteriores ao sítio que irá tornar-se a ponta 3' do transcrito final. Isto é, a transcrição continua além do sítio que irá tornar-se a ponta 3', e o segmento distal é removido por clivagem endonucleolítica. O evento de clivagem que produz a ponta 3' de um transcrito geralmente ocorre no sítio de 11 a 30 nucleotídeos posteriores à sequência conservada de consenso AAUAAA e antecedente à sequência rica em G-U situada perto da ponta da unidade de transcrição. Após a clivagem, a enzima poli(A)-polimerase

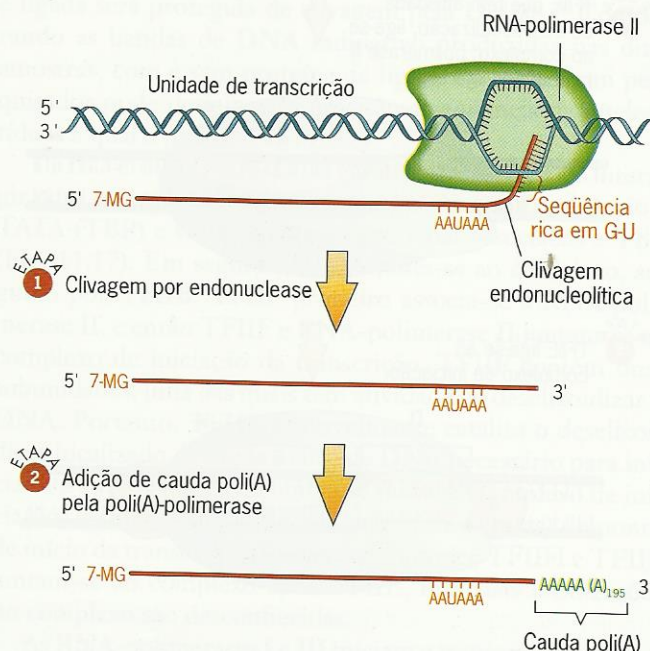


Fig. 11.20 ■ Caudas poli(A) são adicionadas às pontas 3' dos transcritos pela enzima poli(A)-polimerase. As pontas 3', substratos para a poli(A)-polimerase, são produzidas por clivagem endonucleolítica do transcrito posterior ao sinal de poliadenilação, que tem a sequência de consenso AAUAAA.

adiciona caudas poli(A), trechos de adenosina-monofosfato com cerca de 200 nucleotídeos de tamanho, às pontas 3' dos transcritos (Fig. 11.20). A adição de caudas poli(A) a mRNA eucarióticos é chamada **poliadenilação**.

A formação de caudas poli(A) em transcritos requer um componente específico que reconhece e se liga à sequência AAUAAA, um fator estimulatório que se liga à sequência rica em G-U, uma endonuclease, e à polimerase poli(A). Estas proteínas formam um complexo multimérico que efetua tanto a clivagem quanto a poliadenilação em reações muito acopladas. As caudas poli(A) de mRNA eucarióticos acentuam sua estabilidade e têm um papel importante em seu transporte do núcleo para o citoplasma.

Em contraste com a RNA-polimerase II, tanto a RNA-polimerase I quanto a III respondem a discretos sinais de término. A RNA-polimerase I termina a transcrição em resposta a uma sequência de 18 nucleotídeos de tamanho que é reconhecida por uma proteína de término associada. A RNA-polimerase III responde a um sinal de término que é similar ao finalizador independente de rô em *E. coli* (veja Fig. 11.13).

Edição do RNA: Alterando o Conteúdo da Informação de Moléculas de mRNA

De acordo com o dogma central da biologia molecular, a informação genética flui do DNA para o RNA para a proteína durante a expressão gênica. Normalmente, a informação genética não é alterada no mRNA intermediário. Entretanto, a descoberta da **edição do RNA** mostrou que ocorrem exceções. Os processos de edição do RNA alteram o conteúdo da informação dos transcritos gênicos de dois modos: (1) mudando as estruturas de bases individuais e (2) inserindo ou deletando monofosfatos de uridina.

O primeiro tipo de edição de RNA, que resulta na substituição de uma base por outra, é raro. Este tipo de edição foi descoberto em estudos dos genes da apolipoproteína B (*apo-B*) e mRNA em coelhos e humanos. As apolipoproteínas são proteínas sanguíneas que transportam alguns tipos de moléculas de gordura no sistema circulatório. No fígado, o mRNA da *apo-B* codifica uma grande proteína com 4.563 aminoácidos de tamanho. No intestino, o mRNA da *apo-B* dirige a síntese de uma proteína com apenas 2.153 aminoácidos de tamanho. Aqui, uma citosina no pré-mRNA é convertida em U, gerando um códon interno UAA de término de tradução, que resulta na apolipoproteína truncada (Fig. 11.21). UAA é um dos três códons que terminam as cadeias polipeptídicas durante a tradução. Se um códon UAA é produzido dentro da região codificante de um mRNA, ele prematuramente termina o polipeptídeo durante a tradução, gerando um produto gênico incompleto. A conversão C → U é catalisada por uma proteína de ligação ao RNA específica de sequência com uma atividade que remove grupos amino das citosinas. Um exemplo similar de edição de RNA foi documentado para um mRNA que especifica uma prote-

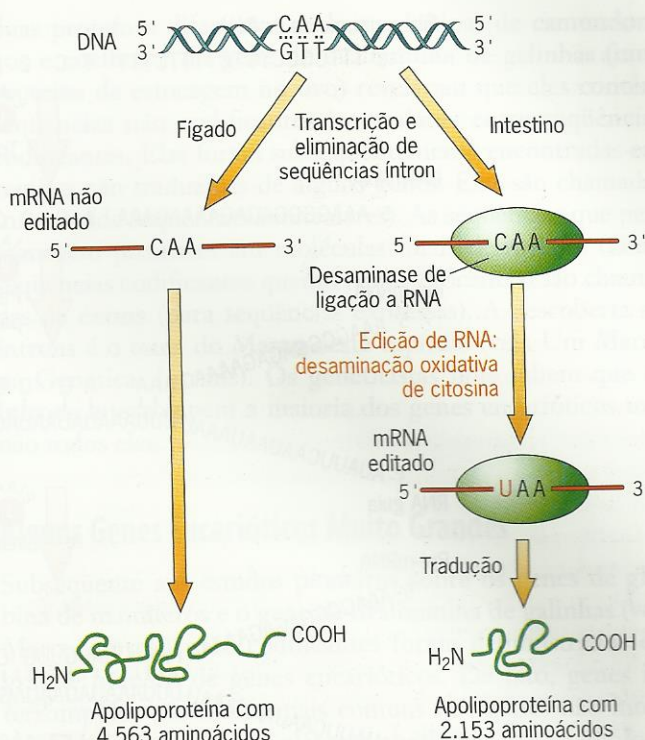


Fig. 11.21 ■ Edição do mRNA da apolipoproteína-B no intestino de mamíferos.

ína (o receptor de glutamato) presente em células do cérebro de rato. A edição mais ampla do mRNA do tipo C → U ocorre nas mitocôndrias das plantas, enquanto a maioria dos transcritos gênicos são editados em algum grau. As mitocôndrias têm seus próprios genomas de DNA e maquinaria de síntese de proteínas (Cap. 19). Em alguns transcritos presentes em mitocôndrias de plantas, a maioria das C são convertidas em U.

Um segundo tipo, mais complexo, de edição de RNA ocorre nas mitocôndrias de tripanossomas (um grupo de protozoários flagelados que causam a doença do sono em humanos). Neste caso, monofosfatos de uridina são inseridos em (e ocasionalmente deletados de) transcritos gênicos, causando grandes mudanças nos polipeptídeos especificados pelas moléculas de mRNA. O tipo de edição mitocondrial de tripanossoma é ilustrado na Fig. 11.22, que mostra as seqüências do filamento codificante de DNA, o RNA transcrito pré-edição e o mRNA editado do gene de citocromo b de *Leishmania tarentolae*. (Citocromo b é uma proteína envolvida na produção de energia química.) Este processo de edição de RNA é mediado por **RNA guias** transcritos de diferentes genes mitocondriais. Os RNA guias contêm seqüências que são parcialmente complementares aos pré-mRNA a serem editados. O pareamento entre os RNA guias e os pré-mRNA resulta em espaços com unidades A não pareadas nos RNA guias. Os RNA guias servem como moldes para edição, e as U são inseridas nos espaços nas moléculas de pré-mRNA em oposição às A nos RNA guias (Fig. 11.22).



Fig. 11.22 ■ Edição do pré-mRNA do citocromo b mitocondrial no tripanossoma *Leishmania tarentolae*. As uridina-monofosfato (U) que foram inseridas nos espaços no pré-mRNA durante o processo de edição estão destacadas em laranja. Pareamento de bases entre o pré-mRNA e o RNA guia é representado por linhas verticais entre as duas moléculas.

Em alguns casos, dois ou mais RNA guias diferentes participam da edição de um único pré-mRNA. Os RNA guias não só fornecem os moldes para os processos de edição, mas também doam monofosfatos de uridina que são inseridos nos pré-mRNA. Os RNA guias têm caudas oligo(U) em 3', com 5 a 24 nucleotídeos de tamanho. Estas U são inseridas nos pré-mRNA durante o processo de edição por um complexo mecanismo de duas etapas (Fig. 11.23). Na primeira etapa, a 3'-OH livre no RNA guia ataca e rompe uma ligação éster interna no pré-mRNA, formando uma ligação covalente com a parte 3' do pré-mRNA clivado. Na segunda etapa, a 3'-OH livre na parte 5' do pré-mRNA clivado ataca uma ligação éster entre duas U nas pontas da cauda oligo(U) do RNA guia, reestruturando a molécula de pré-mRNA, mas agora com uma U inserida no sítio do ataque inicial.

Por que ocorrem estes processos de edição de RNA? Por que as seqüências finais de nucleotídeos destes mRNA não são especificadas pelas seqüências dos genes mitocondriais como são na maioria dos genes nucleares? As respostas a estas interessantes perguntas são puramente especulativas. Os tripanossomas são eucariontes unicelulares primitivos que

divergiram de outros eucariontes bem cedo na evolução. Alguns evolucionistas especulam que a edição do RNA era comum em células ancestrais, em que muitas reações parecem ser catalisadas por moléculas de RNA em vez de proteínas. Outra posição é que a edição do RNA é um mecanismo primitivo para alterar os padrões de expressão gênica. Por qualquer motivo, a edição do RNA tem um papel importante na expressão de genes nas mitocôndrias de tripanossomas e plantas.

■ PONTOS IMPORTANTES

- Três RNA-polimerases diferentes estão presentes em eucariontes, e cada polimerase transcreve um conjunto distinto de genes.
- Os transcritos gênicos eucarióticos geralmente sofrem três modificações importantes: (1) a adição de caps de 7-metilguanossina às pontas 5', (2) a adição de caudas poli(A) na ponta 3' e (3) a remoção de seqüências íntron não codificantes.
- O conteúdo de informações de alguns transcritos eucarióticos é alterado pela edição do RNA, que muda a seqüência de nucleotídeos dos transcritos antes de sua tradução.

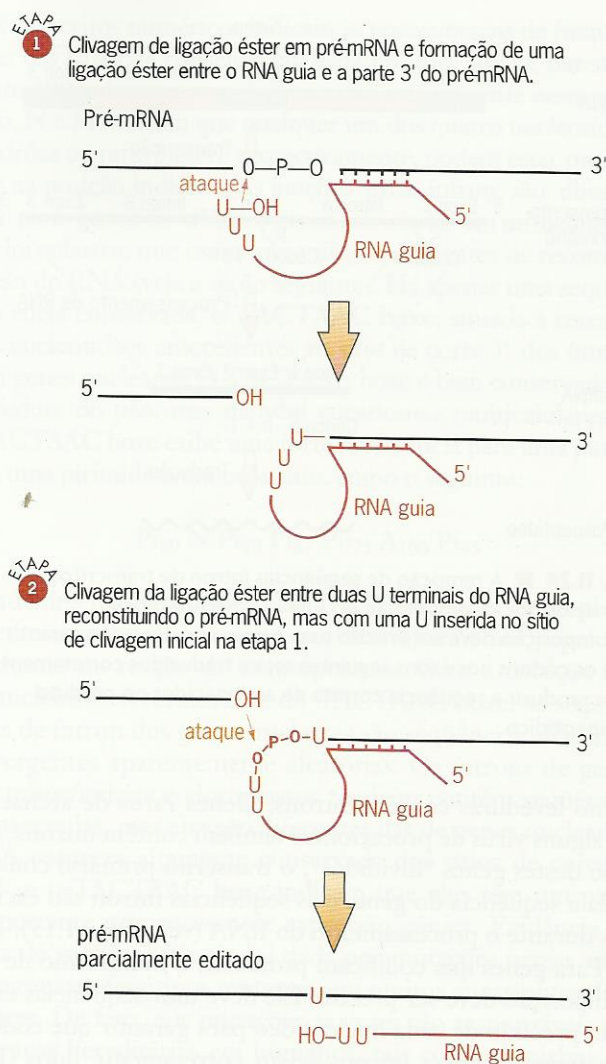


Fig. 11.23 ■ O mecanismo de duas etapas pelo qual uridina-monofosfato são inseridas em moléculas de pré-mRNA durante sua edição em mitocôndrias de tripanossomas.

GENES INTERROMPIDOS EM EUKARIOTES: ÉXONS E ÍNTRONS

A maioria dos genes eucarióticos contém seqüências não codificantes chamadas íntrons que interrompem as seqüências codificantes, os éxons. Os íntrons são removidos dos RNA transcritos antes de seu transporte para o citoplasma.

A maioria dos genes bem caracterizados de procariontes consistem em seqüências contínuas de pares de nucleotídeos, que especificam seqüências co-lineares de aminoácidos nos produtos gênicos polipeptídicos. Entretanto, em 1977, a análise molecular de três genes eucarióticos produziu uma grande surpresa. Estudos de genes de β -globina (uma das

duas proteínas diferentes na hemoglobina) de camundongos e coelhos e do gene de ovalbumina de galinhas (uma proteína de estocagem no ovo) revelaram que eles contêm seqüências não codificantes intercalares entre seqüências codificantes. Elas foram subsequentemente encontradas em regiões não traduzidas de alguns genes. Elas são chamadas **íntrons** (de seqüências **intercalares**). As seqüências que permanecem presentes em moléculas finais de mRNA (tanto seqüências codificantes quanto não codificantes) são chamadas de **éxons** (para seqüências **expressas**). A descoberta de íntrons é o tema do Marco deste capítulo (veja Um Marco na Genética: Íntrons). Os geneticistas hoje sabem que os íntrons interrompem a maioria dos genes eucarióticos mas não todos eles.

Alguns Genes Eucarióticos Muito Grandes

Subseqüente aos estudos pioneiros sobre os genes de globina de mamíferos e o gene de ovalbumina de galinhas (veja Marco), íntrons não codificantes foram demonstrados em grande número de genes eucarióticos. De fato, genes interrompidos são muito mais comuns que genes não interrompidos em animais e plantas superiores. Por exemplo, o gene de *Xenopus laevis* que codifica a vitelogenina A2 (que gera a proteína da gema do ovo) contém 33 íntrons, e o gene do colágeno 1 α 2 de galinha contém pelo menos 50 íntrons. O gene do colágeno tem 37.000 pares de nucleotídeos, mas origina uma molécula de mRNA com apenas cerca de 4.600 nucleotídeos de tamanho. Outros genes contêm relativamente poucos íntrons, mas alguns dos íntrons são muito grandes. Por exemplo, o gene *Ultrabithorax* (*Ubx*) de *Drosophila* contém um íntron que possui aproximadamente 70.000 pares de nucleotídeos de tamanho. O maior gene caracterizado hoje em dia é o gene humano *DMD*, que causa a distrofia muscular Duchenne quando tornado não funcional por mutação. O gene *DMD* tem 2,5 milhões de pares de nucleotídeos e contém 78 íntrons.

Embora os íntrons estejam presentes na maioria dos genes de animais e plantas superiores, eles não são essenciais porque nem todos estes genes contêm íntrons. Os genes de histona do ouriço do mar e quatro genes *heat shock* (choque de calor) de *Drosophila* estavam entre os primeiros genes animais que não tinham íntrons. Hoje sabemos que muitos genes de animais e plantas superiores não têm íntrons.

Íntrons: Significado Biológico?

No momento, os cientistas sabem relativamente pouco sobre o significado biológico da estrutura éxon-íntron de genes eucarióticos. Os íntrons são altamente variáveis em tamanho, indo de cerca de 50 pares de nucleotídeos até milhares de pares de nucleotídeos de tamanho. Este fato levou à especulação de que os íntrons possam ter um papel na regulação da expressão gênica. A transcrição de genes com grandes íntrons leva mais tempo do que a transcrição de

genes com pequenos íntrons ou sem íntrons. Assim, seria esperado que a presença de grandes íntrons em genes reduziria a taxa de acúmulo de transcritos em uma célula. O fato de que os íntrons acumulam novas mutações muito mais rapidamente que os éxons indica que as seqüências de pares de nucleotídeos específicas dos íntrons, excluindo as pontas, não são muito importantes.

Em alguns casos, os éxons diferentes dos genes codificam domínios funcionais diferentes dos produtos gênicos protéicos. Isto é mais aparente no caso de genes que codificam cadeias pesadas e leves de anticorpos (Cap. 23). No caso de genes de globina de mamíferos, o éxon do meio codifica o domínio de ligação de hemo da proteína. Tem havido uma considerável especulação de que a estrutura éxon-íntron de genes eucarióticos tenha resultado da evolução de novos genes pela fusão de genes ancestrais não interrompidos (único éxon). Se esta hipótese estiver correta, os íntrons podem ser meramente relíquias do processo evolutivo.

Alternativamente, os íntrons podem fornecer uma vantagem seletiva aumentando as taxas em que as seqüências codificantes em éxons diferentes de um gene podem reestruturar-se por recombinação, acelerando assim o processo de evolução. Em alguns casos, vias alternativas de recomposição de um transcrito produzem uma família de proteínas correlatas (veja Fig. 14.27). Nesses casos, os íntrons resultam em vários produtos de um único gene. A recomposição alternativa do transcrito troponina T de rato é ilustrada na Fig. 21.4. No caso do gene mitocondrial de levedura que codifica o citocromo *b*, os íntrons contêm éxons de genes que codificam enzimas envolvidas no processamento do transcrito primário do gene. Assim, íntrons diferentes podem de fato ter papéis diferentes, e muitos íntrons podem não ter significado biológico. Como muitos genes eucarióticos não têm íntrons, estas regiões não codificantes não são necessárias para a expressão gênica normal.

PONTOS IMPORTANTES

- A maioria dos genes eucarióticos, mas não todos eles, são divididos em seqüências codificantes chamadas éxons e não codificantes chamadas íntrons.
- Alguns genes contêm íntrons muito grandes; outros abrigam grandes números de pequenos íntrons.
- O significado biológico dos íntrons ainda está aberto a debates.

REMOÇÃO DE ÍNTRONS POR RECOMPOSIÇÃO DO RNA

Os íntrons não codificantes são removidos de transcritos gênicos por vários mecanismos diferentes.

A maioria dos genes nucleares que codificam proteínas em eucariontes multicelulares contêm íntrons. Menos, mas ainda muitos, dos genes de eucariontes unicelulares tais

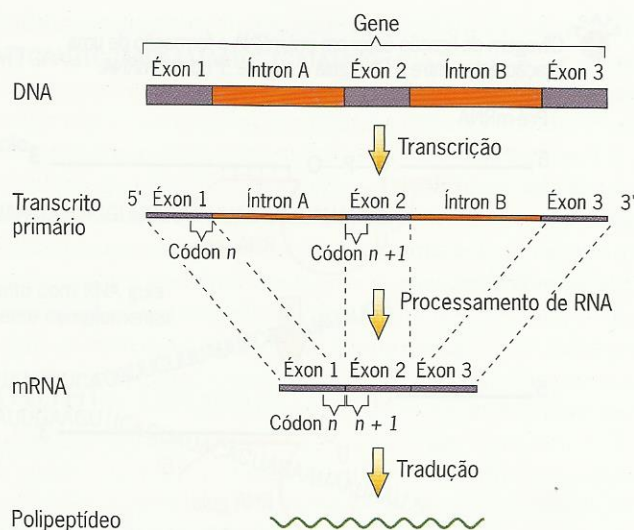
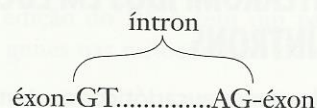


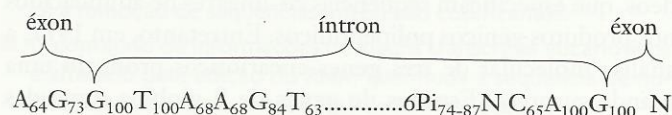
Fig. 11.24 ■ A remoção de seqüências íntron de transcritos primários por recomposição do RNA. O mecanismo de recomposição deve ser preciso para o único nucleotídeo garantir que os códons nos éxons seguintes sejam traduzidos corretamente para produzir a seqüência correta de aminoácidos no produto polipeptídico.

como leveduras contêm íntrons. Genes raros de archaea e de alguns vírus de procariontes também contêm íntrons. No caso destes genes “divididos”, o transcrito primário contém toda a seqüência do gene, e as seqüências íntron são excisadas durante o processamento do RNA (veja Fig. 11.15).

Para genes que codificam proteínas, o mecanismo de recomposição deve ser preciso. Ele deve unir seqüências éxon com precisão de cada nucleotídeo para garantir que códons em éxons distais a íntrons sejam corretamente lidos (Fig. 11.24). A precisão a este grau parece requerer sinais precisos de recomposição, supostamente seqüências de nucleotídeos dentro de íntrons e nas junções éxon-íntron. Entretanto, nos transcritos primários de genes nucleares, as únicas seqüências completamente conservadas de íntrons diferentes são seqüências de dinucleotídeos nas pontas dos íntrons, como por exemplo,



As seqüências mostradas aqui são para o filamento de DNA não-molde (equivalente ao RNA transcrito, mas com T em lugar de U). Além disso, existem curtas seqüências de consenso nas junções éxon-íntron. Para genes nucleares, as junções de consenso são



Os subscritos numéricos indicam as porcentagens de frequências das bases de consenso em cada posição. Assim, um subscrito 100 indica que uma base sempre está presente nesta posição. N e Pi indicam que qualquer um dos quatro nucleotídeos padrões ou pirimidinas, respectivamente, podem estar presentes na posição indicada. As junções éxon-íntron são diferentes para genes de tRNA e genes estruturais em mitocôndrias e cloroplastos, que usam mecanismos diferentes de recomposição do RNA (veja a seção seguinte). Há apenas uma sequência curta conservada, o **TACTAAC box**, situada a cerca de 30 nucleotídeos antecedentes ao sítio de corte 3' dos íntrons em genes nucleares. O TACTAAC box é bem conservado na levedura do pão, mas não em eucariontes multicelulares. O TACTAAC box exibe uma forte preferência para uma purina ou uma pirimidina em cada sítio, como o seguinte:



A adenina na sexta posição do TACTAAC box é completamente conservada e se sabe que desempenha um papel importante na reação de recomposição. Com a exceção dos dinucleotídeos terminais e do TACTAAC box, as sequências de íntron dos genes nucleares são sequências altamente divergentes aparentemente aleatórias. Os íntrons de genes de mitocôndrias e cloroplastos também contêm sequências conservadas, mas elas são diferentes das de genes nucleares.

A natureza altamente conservada dos sítios de corte 5' e 3' e o TACTAAC box indicam que eles têm um papel importante no processo de expressão gênica. Evidência direta de sua importância foi dada por mutações nesses sítios que causam fenótipos mutantes em muitos eucariontes diferentes. De fato, tais mutações às vezes são responsáveis por doenças hereditárias em humanos, tais como distúrbios da hemoglobina.

A descoberta de íntrons não codificantes em genes estimulou o interesse nos mecanismos pelos quais sequências íntron são removidas durante a expressão gênica. A demonstração inicial de que as sequências íntron em genes eucarióticos eram transcritas juntamente com as sequências éxon enfocou pesquisas no processamento de transcritos gênicos primários. Do mesmo modo que os sistemas *in vitro* deram

informações importantes sobre os mecanismos de transcrição e tradução, a chave para a compreensão de eventos de recomposição do RNA foi o desenvolvimento de sistemas de recomposição *in vitro*. Usando estes sistemas, pesquisadores mostraram que existem três tipos distintos de remoção de íntrons dos RNA transcritos.

1. Os íntrons de precursores de tRNA são removidos por clivagem endonucleolítica precisa e reações de ligação catalisadas por endonuclease especial de recomposição e atividades de ligase.
2. Os íntrons de alguns precursores de rRNA são removidos autocataliticamente em uma única reação mediada pela própria molécula de RNA. (Não está envolvida nenhuma atividade enzimática protéica.)
3. Os íntrons de transcritos nucleares pré-mRNA (hnRNA) são removidos em reações de duas etapas efetuadas por complexas partículas de ribonucleoproteínas chamadas spliceossomos.

Estes três mecanismos de remoção de íntrons são discutidos nas três seções seguintes. Existem outros mecanismos de remoção de íntron, mas por questão de simplificação eles não serão discutidos aqui.

Recomposição de Precursor de tRNA: Atividades Únicas de Nuclease e Ligase

A reação de recomposição do precursor de tRNA foi esclarecida em detalhes na levedura *Saccharomyces cerevisiae*. Tanto os sistemas de recomposição *in vitro* quanto os mutantes de recomposição sensíveis a temperatura foram usados para dissecar o mecanismo de recomposição de tRNA em *S. cerevisiae*. A remoção de íntrons de precursores de tRNA em leveduras ocorre em dois estágios (Fig. 11.25). No estágio I, uma **endonuclease de recomposição** nuclear ligada à membrana faz dois cortes exatamente nas pontas do íntron. Então, no estágio II, uma **ligase de recomposição** junta as duas metades do tRNA para produzir a forma final da molécula de tRNA. A especificidade destas reações reside nas

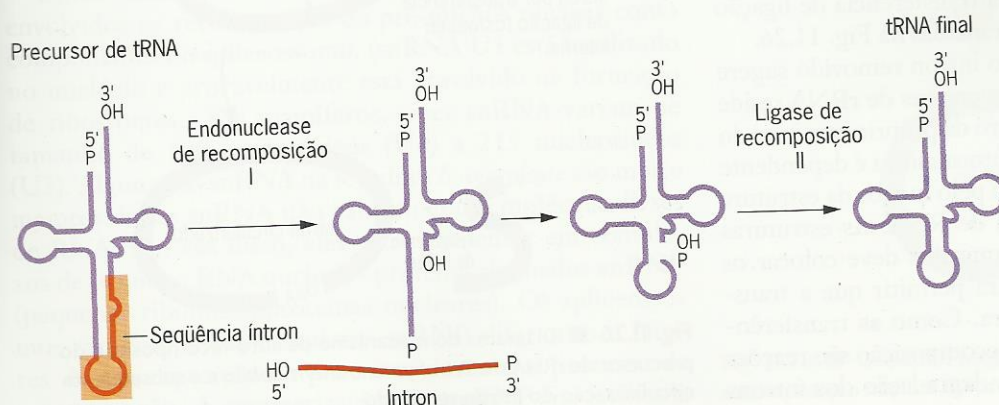


Fig. 11.25 ■ A remoção de íntrons dos precursores de tRNA é um processo de dois estágios: (I) clivagens endonucleolíticas nas extremidades do íntron e (II) ligação das pontas recém-formadas dos dois éxons. O processo de ligação envolve três reações distintas, todas catalisadas pela ligase de recomposição. Esta enzima única contém atividades de fosfodiesterase, cinase e ligase.

características tridimensionais conservadas dos precursores de tRNA, e não nas seqüências de nucleotídeos em si.

Recomposição Autocatalítica

Um tema geral na biologia é que o metabolismo ocorre mediante seqüências de reações catalisadas por enzimas. Essas enzimas muito importantes em geral são proteínas, às vezes polipeptídeos únicos e às vezes complexos heteromultímeros. Ocasionalmente, as enzimas precisam de co-fatores não protéicos para desempenhar suas funções. Quando ligações covalentes estão sendo alteradas, geralmente é suposto que a reação está sendo catalisada por uma enzima. Assim, a descoberta em *Tetrahymena thermophila* de que o íntron no precursor de rRNA foi removido sem o envolvimento de nenhuma atividade catalítica protéica surpreendeu. Entretanto, hoje está claramente estabelecido que a atividade de recomposição que remove o íntron deste precursor de rRNA é intrínseca à própria molécula de RNA. Além disso, foi demonstrado que tal **auto-recomposição** ou **atividade autocatalítica** ocorre em precursores de rRNA de vários eucariontes inferiores e em grande número de precursores de rRNA, tRNA e mRNA em mitocôndrias e cloroplastos de muitas espécies diferentes. No caso de muitos destes íntrons, o mecanismo de auto-recomposição é o mesmo daquele usado pelos precursores de rRNA de *Tetrahymena* ou muito semelhante a ele (veja Fig. 11.26). Para outros, o mecanismo de auto-recomposição é similar ao mecanismo de recomposição observado com precursores nucleares de mRNA, mas sem o envolvimento do spliceossomo (veja Fig. 11.27).

A excisão autocatalítica do íntron no precursor de rRNA de *Tetrahymena* e de alguns outros íntrons não requer uma fonte de energia externa e nenhuma atividade catalítica protéica. Em vez disso, o mecanismo de recomposição envolve uma série de transferências de ligações fosfoéster, sem ligações perdidas ou ganhas no processo. A reação requer um nucleosídeo ou nucleotídeo guanina com um grupo 3'-OH livre (GTP, GDP, GMP ou guanosina funcionais) como co-fator mais um cátion monovalente e um cátion divalente. A necessidade de G-3'-OH é absoluta. Nenhuma outra base pode ser substituída no co-fator nucleosídeo ou nucleotídeo. O íntron é removido por meio de duas transferências de ligação fosfoéster, e o íntron removido pode subsequente ser circularizado por meio de outra transferência de ligação fosfoéster. Estas reações estão diagramadas na Fig. 11.26.

A circularização autocatalítica do íntron removido sugere que a auto-recomposição destes precursores de rRNA reside primariamente, ou totalmente, dentro da própria estrutura do íntron. Supostamente, a atividade autocatalítica é dependente da estrutura secundária do íntron ou pelo menos da estrutura secundária da molécula precursora de RNA. As estruturas secundárias desses RNA auto-recompostos deve colocar os grupos reativos em justaposição para permitir que a transferência da ligação fosfoéster ocorra. Como as transferências das ligações fosfoéster de auto-recomposição são reações potencialmente reversíveis, a rápida degradação dos íntrons

removidos, ou o envio de rRNA recompostos para o citoplasma, pode ativar a recomposição no sentido para diante.

Note que as reações de recomposição autocatalíticas são de natureza intramolecular e portanto não são dependentes de concentração. Além disso, os precursores de RNA são capazes de formar um centro ativo ao qual se liga o co-fator guanosina-3'-OH. A recomposição autocatalítica desses precursores de rRNA demonstra que os sítios catalíticos não são restritos a proteínas. Entretanto, não há atividade *trans* catalítica como para enzimas, mas apenas atividade *cis* catalítica. Alguns cientistas acreditam que a recomposição

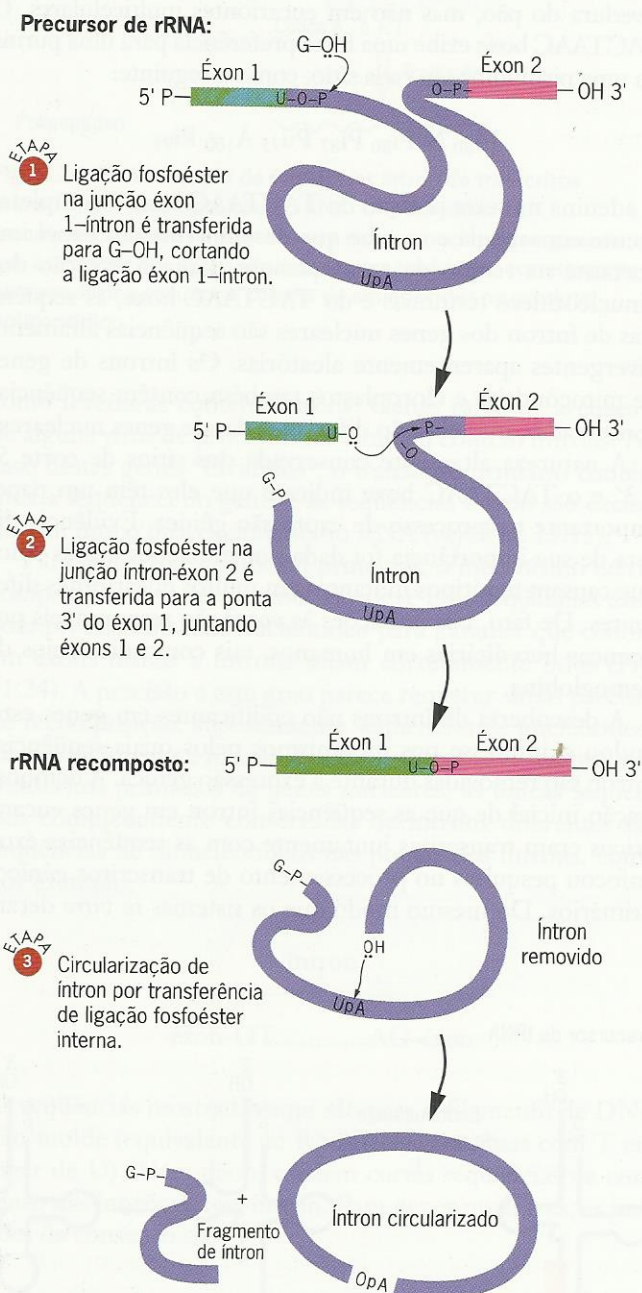


Fig. 11.26 ■ Diagrama do mecanismo de auto-recomposição do precursor de rRNA de *Tetrahymena thermophila* e a subsequente circularização do íntron removido.

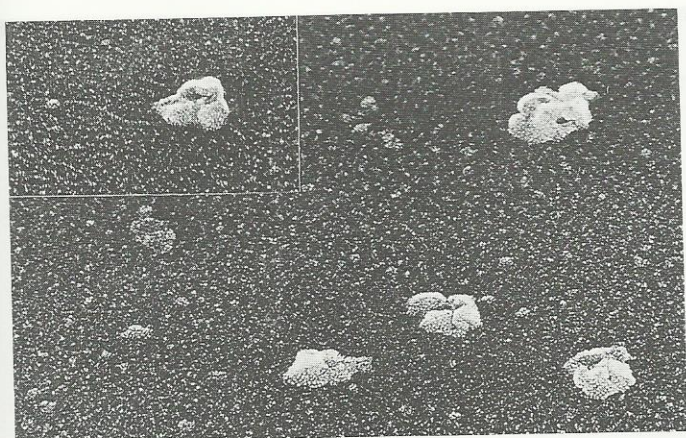


Fig. 11.27 ■ Fotos de microscopia eletrônica de spliceossomos purificados. Note a marcante subestrutura das partículas, que têm dimensões de 40 a 60 nanômetros. O detalhe (acima à esquerda) mostra uma partícula com um filamento fino e uma partícula menor no final do filamento.

autocatalítica do RNA possa ser uma relíquia de um mundo inicialmente baseado em RNA.

Recomposição de Pré-mRNA: snRNA, snRNP e o Spliceossomo

Os íntrons em pré-mRNAs nucleares são removidos em duas etapas, como os íntrons nos precursores de tRNA de leveduras e precursores de rRNA de *Tetrahymena* que foram discutidos nas duas seções precedentes. Entretanto, os íntrons não são removidos por simples nucleases e ligases de recomposição ou autocataliticamente, e não é necessário nenhum co-fator guanosina. Em vez disso, a recomposição de pré-mRNA nuclear é feita por estruturas complexas de RNA/proteína chamadas **spliceossomos** (Fig. 11.27). Estas estruturas são de muitos modos como pequenos ribossomos. Elas contêm um conjunto de pequenas moléculas de RNA denominadas snRNA (RNA pequenos nucleares) e cerca de 40 proteínas diferentes. As duas etapas na recomposição do pré-mRNA nuclear são conhecidas (Fig. 11.28). Entretanto, alguns detalhes do processo de recomposição ainda são incertos.

Cinco snRNA, chamados U1, U2, U4, U5 e U6, estão envolvidos na recomposição do pré-mRNA nuclear como componentes do spliceossomo. (snRNA U3 está localizado no nucléolo e provavelmente está envolvido na formação de ribossomos.) Em mamíferos, estes snRNA variam de tamanho de 100 nucleotídeos (U6) a 215 nucleotídeos (U3). Alguns dos snRNA na levedura *S. cerevisiae* são muito maiores. Esses snRNA não existem como moléculas livres de RNA. Em vez disso, eles estão presentes em complexos de pequeno RNA nuclear – proteína chamados **snRNP** (pequenas ribonucleoproteínas nucleares). Os spliceossomos são compostos de quatro snRNP diferentes e fatores de recomposição de proteínas durante o processo de recomposição. A caracterização dos snRNP foi facilitada

pela descoberta de que alguns pacientes com uma doença auto-imune às vezes fatal chamada lúpus eritematoso sistêmico produzem anticorpos que reagem com muitos de seus próprios componentes celulares incluindo proteínas snRNP. Tais anticorpos são chamados *auto-anticorpos* porque eles reagem com as proteínas do próprio paciente. Normalmente, o sistema imune humano produzirá apenas anticorpos que reagem com proteínas exógenas (Cap. 23). Em pacientes com lúpus, os auto-anticorpos causam inflamações de tecidos e órgãos que podem resultar em funcionamento anormal do coração, dos rins ou do fígado e até mesmo na morte. Os auto-anticorpos de pacientes com lúpus eritematoso sistêmico podem ser usados para precipitar snRNP. Assim, eles facilitam muito a purificação dos snRNP para estudos estruturais e funcionais.

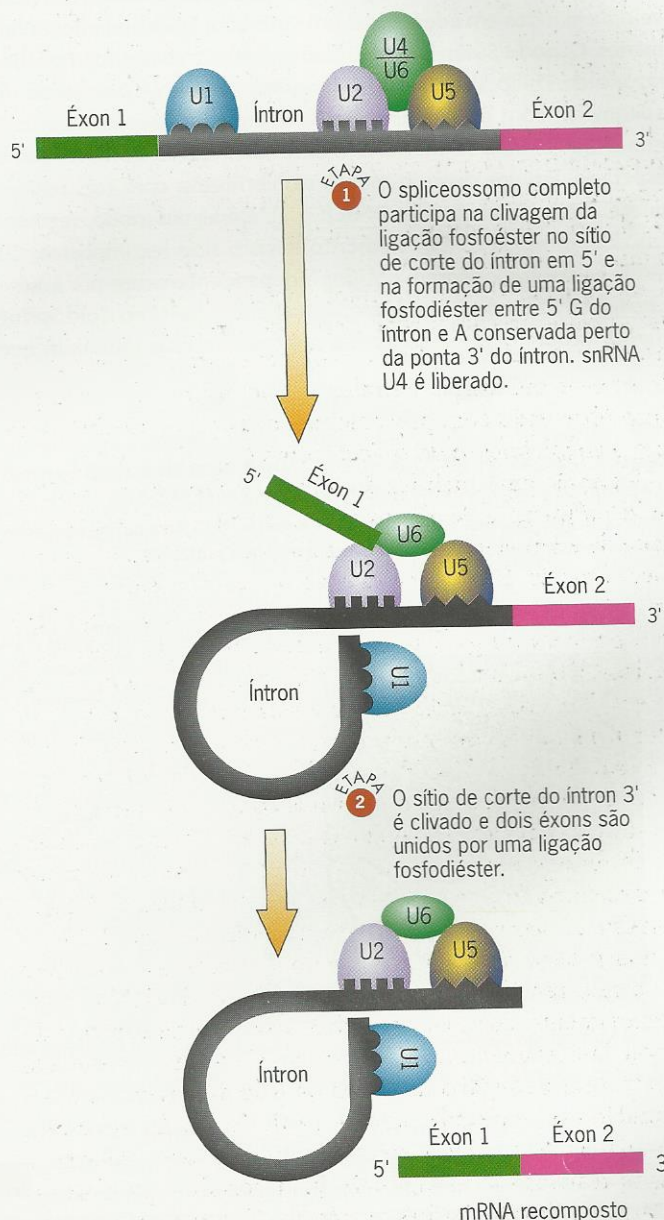


Fig. 11.28 ■ Os papéis postulados dos snRNP contendo snRNA na recomposição nuclear do pré-mRNA.



UM MARCO NA GENÉTICA: Introns

Durante os anos 1960 e o início dos anos 1970, surgiu uma visão elegantemente simples do gene. O gene era uma sequência linear de pares de bases codificando uma sequência co-linear de aminoácidos em seus produtos polipeptídicos. Esta visão do gene foi desenvolvida quase que exclusivamente mediante estudos de procariontes e seus vírus. Logicamente, nós também sabemos que apenas 1 a 2 por cento do DNA em mamíferos especificou as sequências de aminoácidos em proteínas. Entretanto, todos foram surpreendidos em 1977 quando sequências não codificantes foram descobertas entre as sequências codificantes de genes eucarióticos.

A primeira evidência de que algo inesperado estava ocorrendo veio de estudos em adenovírus, um vírus com DNA que infecta humanos. Quando Susan Berget e Claire Moore, trabalhando no laboratório de Philip Sharp no MIT, hibridizaram mRNA purificado de adenovírus com DNA adenoviral unifilamentar, eles descobriram que o mRNA não se hibridizava com uma região do DNA como esperado. Em vez disso, o mRNA se hibridizava com três regiões diferentes no filamento de DNA (Fig. 1).¹ De algum modo, segmentos diferentes do mRNA transcrito haviam sido recompostos. Os mesmos resultados foram obtidos independentemente por Louise Chow e colaboradores no grupo de Richard Roberts no Cold Spring Harbor Laboratory.² Nessa época, muitos cientistas pensaram que

tal estranha recomposição de segmentos do mRNA podia ser algo único de vírus. Entretanto, essa visão logo se mostrou errada.

Pesquisas em quatro laboratórios diferentes demonstraram que sequências intercalares estavam presentes em genes codificantes de tipos diferentes de proteínas em vertebrados. O trabalho sobre o gene de β -globina de coelho no laboratório de Richard Flavell,³ o gene de β -globina de comundongo no laboratório de Philip Leder⁴ e o gene de ovalbumina de galinha nos laboratórios de Pierre Chambon⁵ e de Bert W. O'Malley⁶ documentaram independentemente a presença de sequências intercalares não codificantes, hoje chamadas introns, nestes genes estruturais. Todos os quatro estudos dependeram do desenvolvimento de métodos pelos quais genes individuais puderam ser isolados ("clonados") e caracterizados (Cap. 15). Quando os genes foram comparados com seus produtos de mRNA, foi visto que eles continham sequências de nucleotídeos que não estavam presentes nos mRNA.

³Jeffreys, A. J., e R. A. Flavell, 1977. The rabbit β -globin gene contains a large insert in the coding sequence. *Cell* 12: 1097-1108.

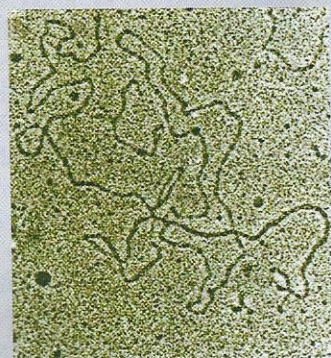
⁴Tilghman, S. M., D. C. Tiemeier, F. Polsky, M. H. Edgell, J. G. Seidman, A. Leder, L. W. Enquist, R. Norman e P. Leder, 1977. Cloning specific segments of the mammalian genome: Bacteriophage λ containing mouse globin and surrounding gene sequences. *Proc. Natl. Acad. Sci. U.S.A.* 74: 4406-4410.

⁵Breathnach, R., C. Benoist, K. O'Hare, F. Gannon e P. Chambon, 1978. Ovalbumin gene: Evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *Proc. Natl. Acad. Sci. U.S.A.* 75: 4853-4857.

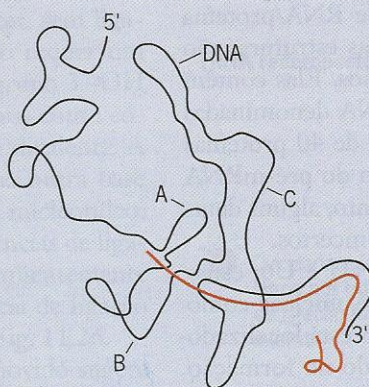
⁶Woo, S. L. C., A. Dugaiczky, M.-J. Tsai, E. C. Lai, J. F. Catterall e B. W. O'Malley, 1978. The ovalbumin gene: Cloning of the natural gene. *Proc. Natl. Acad. Sci. U.S.A.* 75: 3688-3692.

¹Berget, S. M., C. Moore e P. A. Sharp, 1977. Spliced segments at the 5' terminus of adenovirus 2 late RNA. *Proc. Natl. Acad. Sci. U.S.A.* 74: 3171-3175.

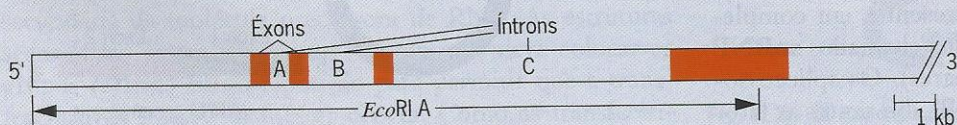
²Chow, L. T., R. E. Gelinas, T. R. Broker e R. J. Roberts, 1977. An amazing sequence arrangement at the 5' ends of adenovirus 2 mRNA. *Cell* 12: 1-9.



(a) Micrografia eletrônica



(b) Diagrama interpretativo de (a).



(c) Localizações das seqüências éxon e intron no genoma de adenovírus.

Fig. 1 ■ A primeira evidência de introns e recomposição do RNA. (a) Micrografia eletrônica mostrando a estrutura do heterodúplex formado por hibridização do mRNA final de adenovírus com o filamento complementar de DNA no cromossomo de adenovírus (reproduzida com permissão de S. M. Berget, C. Moore e P. A. Sharp, 1977. *Proc. Natl. Acad. Sci. U.S.A.* 74: 3171-3175). (b) Desenho interpretativo da estrutura mostrada em (a). (c) Diagrama mostrando as localizações dos introns (seqüências intercalares) e éxons na ponta 5' do cromossomo de adenovírus.



UM MARCO NA GENÉTICA (continuação)

Algumas das primeiras evidências de íntrons em genes de β -globina de mamíferos resultaram da visualização de híbridos genômicos de DNA-mRNA por microscopia eletrônica. Os duplícies DNA-RNA são mais estáveis que a dupla hélice de DNA. Assim, se as duplas hélices de DNA parcialmente desnaturadas são incubadas com moléculas homólogas de RNA em condições apropriadas, os filamentos de RNA irão hibridizar com filamentos complementares de DNA, deslocando os filamentos equivalentes de DNA (Fig. 2). As estruturas híbridas DNA-RNA resultantes conterão regiões unifilamentares de DNA chamadas **alças R**, onde as moléculas de RNA deslocaram filamentos de DNA para formar regiões duplícies DNA-RNA. Essas alças R podem ser visualizadas diretamente por microscopia eletrônica. Assim a hibridização de alça R e a microscopia eletrônica proporcionaram um poderoso instrumento para estudos da estrutura gênica.

Quando Shirley Tilghman, Philip Leder e colaboradores hibridizaram mRNA purificado de β -globina de camundongo com uma molécula de DNA que continha o gene de β -globina de camundongo, eles observaram duas alças R separadas por uma alça de DNA bifila-

mentar (Fig. 3a).⁷ Seus resultados demonstraram a presença de uma sequência de pares de nucleotídeos no meio do gene de β -globina que não está presente no mRNA de β -globina e, portanto, não codifica aminoácidos no polipeptídeo de β -globina. Quando Tilghman e colaboradores repetiram os experimentos com alça R usando transcritos do gene purificado de β -globina isolados de núcleos e tidos como sendo transcritos gênicos primários ou moléculas de pré-mRNA, em lugar do mRNA citoplasmático de β -globina, eles observaram apenas uma alça R (Fig. 3b). Este resultado indicou que o transcrito primário contém a sequência completa do gene estrutural, incluindo tanto éxons quanto íntrons. Juntos, os resultados de alças R obtidos com mRNA citoplasmático e pré-mRNA nuclear demonstram que as sequências íntron são removidas e as sequências éxon são recompostas durante o processamento que converte o transcrito primário no mRNA final.

Tilghman e colaboradores logo verificaram sua interpretação dos resultados de alça R seqüenciando um segmento do gene de β -globina de camundongo que envolve a junção entre este grande íntron e o éxon precedente.⁸ Quando eles usaram os códons estabelecidos para comparar a sequência de nucleotídeos do gene de β -globina de camundongo com a sequência conhecida de aminoácidos da β -globina de camundongo, a presença da sequência não codificante ficou clara (Fig. 4). Se não houver íntron nesta posição no gene de β -globina de camundongo, a sequência de nucleotídeos prevê que os aminoácidos 105 até 110 seriam Val-Ser-Leu-Met-Gli-Tre. Entretanto, a sequência de aminoácidos de β -globina de camundongo foi determinada experimentalmente, e os aminoácidos 105 até 110 são Leu-Leu-Gli-Asn-Met-Ile. Em uma publicação posterior, Tilghman, Leder e colaboradores seqüenciaram o resto do gene de β -globina de camundongo e mostraram que os aminoácidos 105 a 110 são codificados pelos primeiros seis pares de trincas de nucleotídeos no éxon distal a este íntron. O grande íntron do gene de β -globina caracterizado nestes primeiros estudos tem 653 pares de nucleotídeos de tamanho. Logo após a descoberta do grande íntron, Tilghman, Leder e colaboradores descobriram que o gene de β -globina de camundongo continha um segundo íntron menor, com 116 pares de nucleotídeos de tamanho (Fig. 4). Os resultados de Richard Flavell e colaboradores mostraram que a estrutura do gene de β -globina de coelho é muito similar à de camundongo.

Estudos de Tom Maniatis, Richard Flavell e outros mostraram que a estrutura do gene humano de β -globina também é muito similar à do gene de β -globina de camundongo. Ambos contêm dois íntrons que separam três éxons. Cada um dos dois genes humanos de α -globina também contêm dois íntrons. De fato, os genes hu-

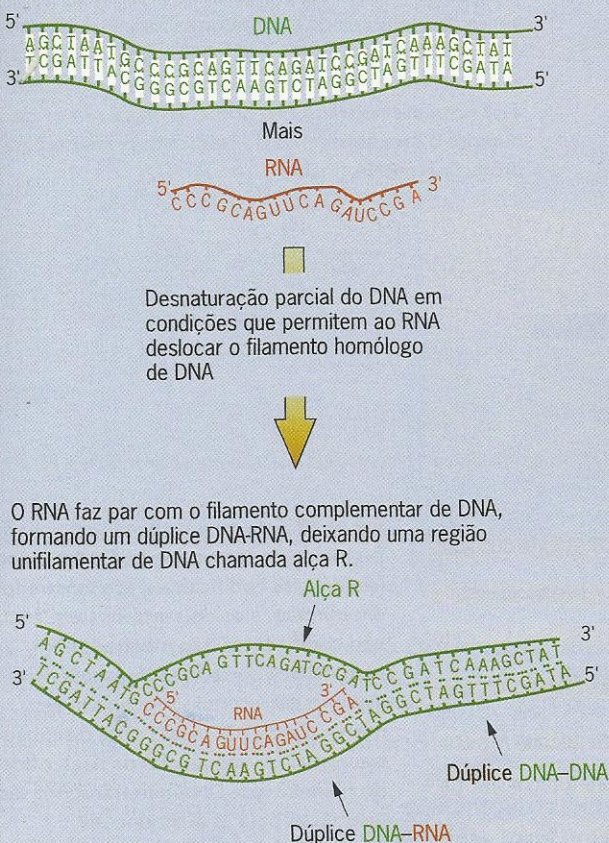


Fig. 2 ■ A técnica de hibridização de alça R.

⁷Tilghman, S. M., D. C. Tiemeier, J. G. Seidman, B. M. Peterlin, M. Sullivan, J. V. Maizel e P. Leder, 1978. Intervening sequence of DNA identified in the structural portion of a mouse β -globin gene. *Proc. Natl. Acad. Sci. U.S.A.* 75: 725-729.

⁸Ibid.



UM MARCO NA GENÉTICA (continuação)

Micrografia eletrônica

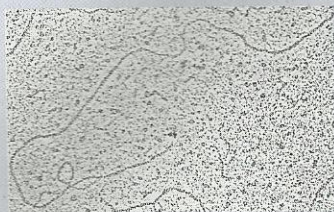
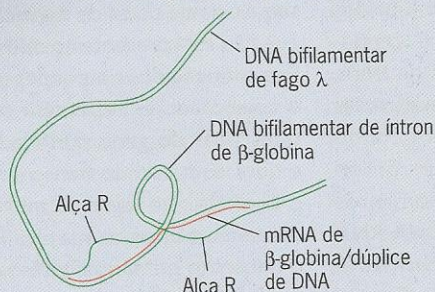


Diagrama interpretativo

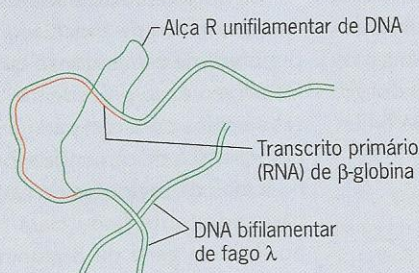


(a) Alças R formadas por mRNA de β-globina.

Micrografia eletrônica



Diagrama interpretativo



(b) Alça R formada pelo transcrito primário (pré-mRNA) de β-globina.

Fig. 3 ■ Evidência de alça R para um intron no gene de β-globina de camundongo. (a) Quando os genes de β-globina de camundongo e os mRNA foram hibridizados em condições de alça R, duas alças R foram observadas nos híbridos RNA-DNA resultantes. (b) Quando transcritos primários de pré-mRNA de genes de β-globina de camundongo foram usados nos experimentos de alça R, observou-se uma única alça R. Estes resultados demonstram que a sequência intron está presente no transcrito primário mas é removida durante o processamento do transcrito primário para produzir o mRNA final.

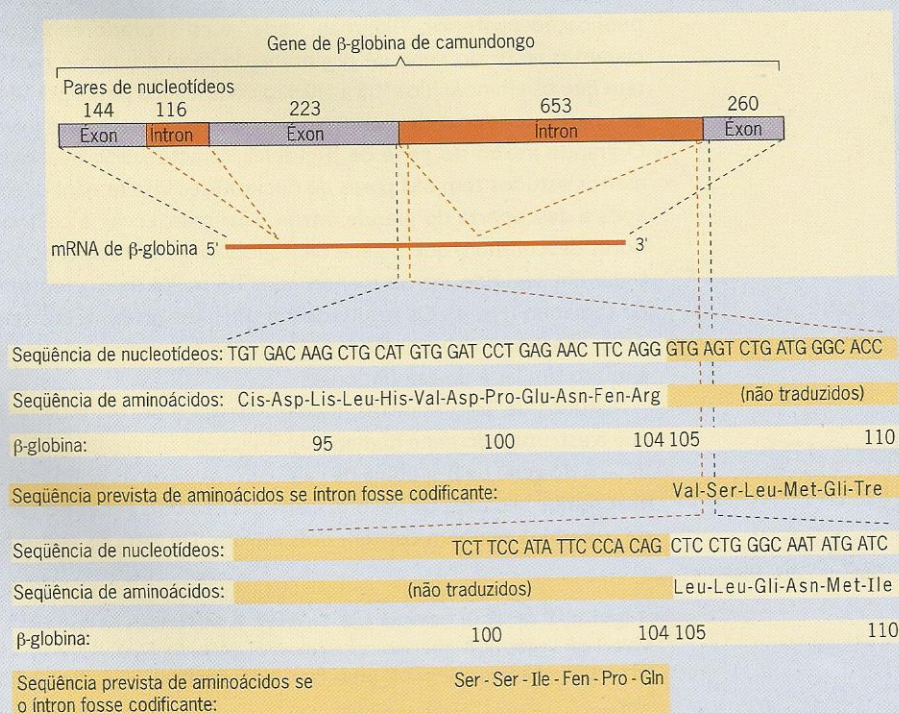


Fig. 4 ■ Estrutura do gene principal de β-globina de camundongo. Os três éxons (seqüências codificantes) são mostrados em púrpura, e os dois introns (seqüências não codificantes) são mostrados em vermelho. A correlação das seqüências de mRNA é indicada pelas linhas diagonais tracejadas. As seqüências de nucleotídeos do filamento de DNA não transcrito são mostradas para as duas junções éxon-intron. Essas seqüências de nucleotídeos são as mesmas das seqüências de mRNA, mas T substitui U.



UM MARCO NA GENÉTICA (continuação)

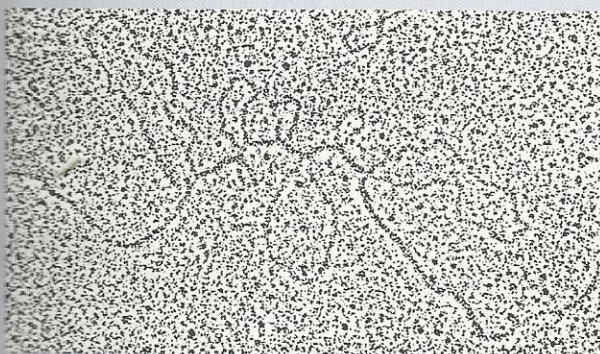
manos codificantes das cadeias de globina embrionária similares a α (ζ) e similares a β (ϵ), da cadeia de globina fetal tipo β (γ) e da cadeia de globina adulta tipo β (δ), todos contêm dois íntrons e três éxons. Além disso, todos os genes humanos que codificam globinas tipo α têm íntrons nas mesmas posições (separando o códon 31 do 32 e o códon 99 do 100), como todos os genes humanos que codificam globinas tipo β (separando o códon 30 do 31 e o códon 104 do 105). Portanto, as posições dos íntrons foram altamente conservadas durante a evolução dos genes de globina humana.

Enquanto Philip Leder, Richard Flavell e colaboradores estavam demonstrando a presença de dois íntrons em genes de β -globina de mamíferos, Pierre Chambon, Bert W. O'Malley e seus colaboradores estavam fazendo experimentos similares no gene de ovalbumina de galinha. Quando o filamento transcrito do gene de ovalbumina de galinha foi hibridizado com o mRNA de ovalbumina e visualizado por microscopia eletrônica, sete alças unifilamentares de DNA foram observadas (Fig. 5), sugerindo que o gene de ovalbumina contém sete íntrons. Experimentos posteriores,

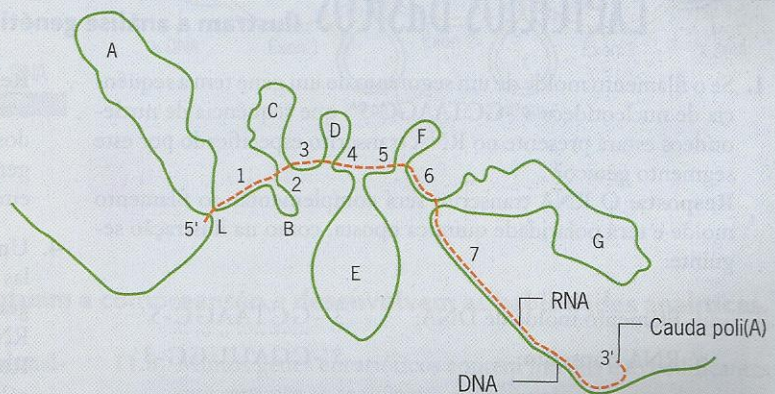
incluindo comparações diretas das seqüências do gene e o mRNA, mostraram conclusivamente que o gene de ovalbumina de galinha contém sete íntrons separando oito éxons. Estudos subsequentes mostraram que os íntrons são uma característica comum da maioria dos genes eucarióticos. Entretanto, ainda temos muito o que aprender sobre suas funções.

QUESTÕES PARA DISCUSSÕES

1. Os íntrons correspondem a quase um quarto do DNA no genoma humano. Os íntrons são apenas seqüências "sucata"? Ou eles têm papéis importantes em organismos eucarióticos? Que funções você pode pensar que tenham as seqüências de DNA situadas em íntrons?
2. Poucos genes procarióticos contêm íntrons, enquanto a maioria dos genes eucarióticos contêm íntrons. Você pode pensar em possíveis explicações para esta diferença? Qual de suas explicações você acha que é mais provavelmente a correta? Por quê?



(a) Micrografia eletrônica do heterodúplce DNA-mRNA de ovalbumina.



(b) Diagrama interpretativo do heterodúplce DNA-mRNA.

Fig. 5 ■ Evidência de sete íntrons no gene de ovalbumina de galinha. A molécula de mRNA é representada por uma linha vermelha tracejada; o filamento de DNA é mostrado como uma linha verde. As seqüências íntron são alças de DNA unifilamentar marcadas de A até G.

Cada um dos snRNA U1, U2 e U5 está presente em uma partícula específica de snRNP. Os snRNA U4 e U6 estão presentes juntos em um quarto snRNP; os snRNA U4 e U6 contêm duas regiões de complementariedade intermolecular que fazem pares de bases no snRNP U4/U6. Cada um dos quatro tipos de partículas de snRNP contém um subgrupo de sete proteínas snRNP bem caracterizadas mais uma ou mais proteínas únicas para o tipo particular de partícula de snRNP. Todos os quatro complexos de snRNP estão presentes nos spliceossomos isolados mostrados na Fig. 11.27.

A primeira etapa na recomposição nuclear de pré-mRNA envolve clivagem no sítio de clivagem do íntron em 5' (\downarrow GU-íntron) e a formação de uma ligação fosfodiéster intramolecular entre o carbono 5' da G no sítio de clivagem e o carbono 2' de uma A conservada perto da ponta 3' do íntron. Esta etapa ocorre em spliceossomos completos (Fig. 11.28) e requer a hidrólise de ATP. Evidências indicam que o snRNP U1 deve ligar-se ao sítio de corte 5' antes da reação inicial de clivagem. O reconhecimento do sítio de clivagem na ponta 5' do íntron provavelmente envolve pareamento de bases

entre a sequência de consenso neste sítio e uma sequência complementar perto da ponta 5' do snRNA U1. Entretanto, a especificidade da ligação de pelo menos alguns dos snRNP às sequências de consenso do íntron envolve tanto os snRNA quanto proteínas snRNP específicas.

A segunda snRNP a ser adicionada ao complexo de recomposição parece ser a snRNP U2. Ela se liga à sequência de consenso que contém a A conservada que forma o ponto de ramificação na estrutura em laço do íntron recomposto. Portanto, o snRNP U5 liga-se ao sítio 3' de recomposição, e o snRNP U4/U6 é adicionado ao complexo para produzir um spliceossomo completo (Figs. 11.27 e 11.28). Quando o sítio de recomposição do íntron 5' é cortado na etapa 1, o snRNA U4 é liberado do spliceossomo. Na etapa 2 da reação de recomposição, o sítio de recomposição em 3' do íntron é cortado, e os dois éxons são unidos por uma ligação fosfodiéster normal 5' com 3' (Fig. 11.28). O mRNA recomposto

está agora pronto para ser mandado para o citoplasma e traduzido em ribossomos.

PONTOS IMPORTANTES

- Sequências íntron não codificantes são removidas dos RNA transcritos no núcleo antes de seu transporte para o citoplasma.
- Os íntrons em precursores de tRNA são removidos pela ação conjunta de uma endonuclease e uma ligase de recomposição enquanto os íntrons em alguns precursores de rRNA são removidos autocataliticamente, sem proteína catalítica envolvida.
- Os íntrons em pré-mRNA nucleares são removidos em estruturas ribonucleoprotéicas complexas chamadas spliceossomos.
- O processo de remoção de íntrons deve ser preciso, com exatidão no nível de nucleotídeo, para garantir que códons em éxons distais a íntrons sejam lidos corretamente durante a tradução.

Exercícios Básicos Ilustram a análise genética básica.

1. Se o filamento molde de um segmento de um gene tem a sequência de nucleotídeos 3'-GCTAAGC-5', que sequência de nucleotídeos estará presente no RNA transcrito especificado por este segmento gênico?

Resposta: O RNA transcrito será complementar ao filamento molde e terá polaridade química oposta, como na ilustração seguinte:

filamento molde de DNA:	3'-GCTAAGC-5'
RNA transcrito:	5'-CGAUUCG-3'

2. Se o filamento não-molde de um gene em *E. coli* tem a sequência:

5'-TTGACA-(18 bases)-TATAAT-(8 bases)-GCCTTCCAGTG-3'

que sequência de nucleotídeos estaria presente no RNA transcrito deste gene?

Resposta: O gene contém sequências promotoras perfeitas -35 e -10. A transcrição deve ser iniciada em um sítio de cinco a nove bases posteriores à sequência -10 TATAAT, e o término 5' do transcrito deve conter uma purina. O filamento molde na ponta 5' do transcrito deve ter a seguinte estrutura:

Filamento molde de DNA:	
(sequência -35)	(sequência -10)
3'-AACTGT-(18 bases)-ATATTA-(8 bases)-CGGAAGGTCAC-5'	
RNA transcrito:	5'-GCCUUCGAGUG-3'

3. Se o filamento não-molde mostrado no exercício 2 fosse parte de um gene em *Drosophila* no lugar de *E. coli*, seria produzido o mesmo transcrito?

Resposta: Não, pois as sequências promotoras que controlam a transcrição em eucariontes tais como *Drosophila* são diferentes dos promotores em procariontes tais como *E. coli*. Portanto, o gene de *E. coli* provavelmente não seria transcrito se presente em *Drosophila*.

4. Uma cientista está estudando a transcrição de um gene em células de camundongo em cultura. Como ela pode determinar se o gene é transcrito por RNA-polimerase I, RNA-polimerase II ou RNA-polimerase III?

Resposta: Ela deve determinar que efeito o tratamento das células com α -amanitina tem na transcrição do gene. Se a α -amanitina não tiver efeito na transcrição, o gene é transcrito pela RNA-polimerase I; se a α -amanitina bloqueia a transcrição, o gene é transcrito pela RNA-polimerase II; se a α -amanitina diminui mas não bloqueia totalmente a transcrição, o gene é transcrito pela RNA-polimerase III.

5. O transcrito primário ou pré-mRNA de um gene nuclear em um chimpanzé tem a sequência:

5'-G—éxon 1—AGGUAAGC—íntron—CAGUC—éxon 2—A-3'

Após o íntron ter sido removido, qual a sequência mais provável do mRNA?

Resposta: Os íntrons contêm pontas dinucleotídicas altamente conservadas: 5'-GT—AG-3' no filamento molde do DNA ou 5'-GU—AG-3' no RNA transcrito. Assim, a sequência íntron é quase certamente 5'-GUUAAGC—íntron—CAG-3'. Com a precisa remoção do íntron, a sequência do mRNA será:

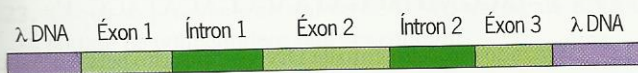
5'-G—éxon 1—AGUC—éxon 2—A-3'

Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. Algumas proteínas humanas de importância médica tais como a insulina e o hormônio de crescimento hoje estão sendo produzidas em bactérias. Usando-se as ferramentas de engenharia genética, seqüências de DNA que codificam estas proteínas foram introduzidas em bactérias. Você quer introduzir um gene humano em *E. coli* e fazer com que este gene produza grandes quantidades do produto gênico humano nas células bacterianas. Supondo-se que o gene humano de interesse possa ser isolado e introduzido em *E. coli*, que problemas você teria em tentar atingir sua meta?

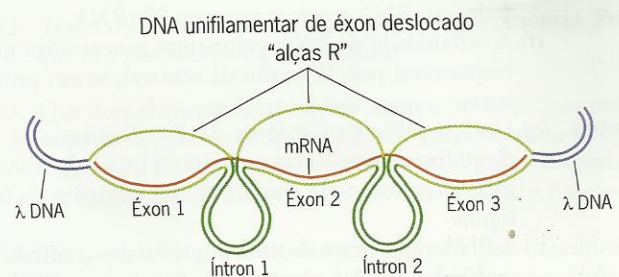
Resposta: As seqüências promotoras necessárias para iniciar a transcrição são muito diferentes em mamíferos e bactérias. Portanto, seu gene não será expresso em *E. coli* a menos que você primeiro junte sua região codificante a uma promotor bacteriano. Além disso, seu gene humano provavelmente conterá íntrons. Como *E. coli* não contém spliceossomos ou maquinaria equivalente que remova íntrons dos RNA transcritos, seu gene humano não será expresso corretamente se contiver íntrons. Como você pode ver, a expressão de genes eucarióticos em células procarióticas não é uma tarefa trivial.

2. Um gene humano de β -globina foi purificado e inserido em um cromossomo linear de bacteriófago lambda, produzindo a seguinte molécula de DNA:



Se esta molécula de DNA é hibridizada com um mRNA humano de β -globina usando-se condições que favorecem dúpiques DNA-RNA em relação a dúpiques DNA-DNA (condições de mapeamento de alça R) e o produto é visualizado por microscopia eletrônica, que estrutura de ácido nucleico você esperaria ver?

Resposta: O transcrito primário deste gene humano de β -globina conterá ambos os íntrons e todos os três éxons. Entretanto, antes de mandá-lo para o citoplasma, as seqüências de íntron serão removidas do transcrito. Assim, a molécula de mRNA final conterá as três seqüências éxon reunidas sem a presença de seqüências íntron. Em condições de alça R, o mRNA irá helicoidizar-se com o filamento complementar de DNA, deslocando o filamento equivalente de DNA. Entretanto, como o mRNA não contém seqüências íntron, os íntrons permanecerão como regiões de DNA bifilamentar como mostrado no diagrama seguinte.



Questões e Problemas Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 11.1. Distinga o DNA do RNA (a) quimicamente, (b) funcionalmente e (c) pela localização na célula.
- 11.2. Que bases no mRNA transcrito representariam a seguinte seqüência molde de DNA: 5'-TGCAGACA-3'?
- 11.3. Que bases no filamento transcrito de DNA dariam origem à seguinte seqüência de bases do mRNA: 5'-CUGAU-3'?
- 11.4. Com base em que evidência foi estabelecida a hipótese do RNA mensageiro?
- 11.5. Em que locais da célula eucariótica ocorre a síntese de proteínas?
- 11.6. Cite três modos pelos quais os mRNA de eucariontes diferem dos mRNA de procariontes.
- 11.7. Que tipos diferentes de moléculas de RNA estão presentes em células procarióticas? em células eucarióticas? Que papéis estas classes diferentes de moléculas de RNA têm na célula?
- 11.8. Muitos genes eucarióticos contêm íntrons não codificantes que separam as seqüências codificantes ou éxons destes genes. Em que estágio durante a expressão destes genes divididos são removidas as seqüências íntron não codificantes?
- 11.9. Por várias décadas, o dogma na biologia tem sido que as reações moleculares em células vivas são catalisadas por enzimas compostas de polipeptídeos. Sabemos hoje que os íntrons de algumas moléculas precursoras de RNA tais como os precursores de rRNA em *Tetrahymena* são removidos autocataliticamente ("auto-recompostos") sem envolvimento de nenhuma proteína catalítica. O que a demonstração da recomposição autocatalítica indica sobre o dogma de que as reações biológicas são sempre catalisadas por enzimas proteináceas?
- 11.10. Que papel(éis) têm os spliceossomos na expressão gênica? Qual a sua estrutura macromolecular?
- 11.11. Que componentes dos íntrons de genes nucleares codificantes de proteínas em eucariontes superiores são conservados e necessários para a remoção correta de seqüências íntron de transcritos primários por spliceossomos?

- 11.12.** Faça a correspondência entre os seguintes termos com cada uma das descrições dadas abaixo. *Termos:* (1) fator sigma (σ); (2) cauda poli(A); (3) TATAAT; (4) éxons; (5) TATAAAA; (6) RNA-polimerase III; (7) íntron; (8) RNA-polimerase II; (9) RNA heterogêneo nuclear (hnRNA); (10) snRNA; (11) RNA-polimerase I; (12) TTGACA; (13) GGCCAATCT (CAAT box).

Descrições:

- Sequência intercalar encontrada em muitos genes eucarióticos.
 - Sequência conservada de nucleotídeos (-30) em promotores eucarióticos envolvida no início da transcrição.
 - Pequenas moléculas de RNA que estão localizadas nos núcleos de células eucarióticas, na maioria como componentes do spliceossomo, que participam na remoção dos íntrons de transcritos de genes nucleares.
 - Uma sequência (-10) no filamento não-molde do promotor de *E. coli* que facilita a deslicoidização localizada de DNA quando em conjunto com a RNA-polimerase.
 - A RNA-polimerase no núcleo que catalisa a síntese de todos os rRNA exceto o pequeno 5S rRNA.
 - A subunidade da RNA-polimerase procariótica que é responsável pela iniciação da transcrição em promotores.
 - Uma sequência promotora de *E. coli* situada 35 nucleotídeos antecedentes ao sítio de início da transcrição. Serve como um sítio de reconhecimento do fator sigma.
 - A RNA-polimerase no núcleo que catalisa a síntese das moléculas de RNA transportador e pequenos RNA nucleares.
 - Uma série de poliadenosinas com 20 a 200 nucleotídeos de tamanho que é adicionada à ponta 3' da maioria dos RNA mensageiros eucarióticos.
 - A RNA-polimerase que transcreve genes nucleares que codificam proteínas.
 - Uma sequência conservada no filamento não-molde de promotores eucarióticos que está localizada a cerca de 80 nucleotídeos antecedentes ao sítio de início de transcrição.
 - Segmentos de um gene eucariótico que correspondem às sequências no RNA transcrito e processado no final do gene.
 - A população de transcritos primários no núcleo de uma célula eucariótica.
- 11.13.** (a) Quais das seguintes sequências de nucleotídeos de pré-mRNA nuclear potencialmente contêm um íntron?
- 5'-UGACCAUGGCGCUAACACUGCCAAUUG-GCAAUACUGACCUGAUAGCAUCAGCCAA-3'
 - 5'-UAGUCUCAUCUGUCCAUUGACUUCGAAACU-GAAUCGUAACUCCUACGUCUAUGGA-3'
 - 5'-UAGCUGUUUGUCAUGACUGACUGGUCAC-UAUCGUACUAACCUGUCAUGCAAUGUC-3'
 - 5'-UAGCAGUUCUGUCGCCUCGUGGUGCU-GCUGGCCUUUCGUCGUCGGGCUUAGCUA-3'

(5) 5'-UAGGUUCGCAUUGACGUACUUCUGAAAC-UACUACUACUACGCAUCGAGUCUCAA-3'

- (b) Um dos cinco pré-mRNA mostrados em (a) pode sofrer recomposição de RNA para remover uma sequência íntron. Que sequência de nucleotídeos do mRNA resultaria deste evento de recomposição?

- 11.14.** Qual a função dos íntrons em genes eucarióticos?

- 11.15.** Um determinado gene é inserido no cromossomo do fago lambda, e ressaltado que contém três íntrons. (a) O transcrito primário deste gene é purificado de núcleos isolados. Quando este transcrito primário é hibridizado em condições de alça R com o cromossomo lambda recombinante que leva o gene, qual será o aspecto das estruturas em alça R? Marque seu diagrama. (b) O mRNA produzido pelo transcrito primário deste gene é então isolado de polirribossomos citoplasmáticos e similarmente examinado pelo procedimento de hibridização de alça R que usa o cromossomo lambda recombinante que leva o gene. Faça um diagrama de como será o aspecto das estruturas em alça R quando é usado o mRNA citoplasmático. Novamente, marque os componentes de seu diagrama.

- 11.16.** Um segmento de DNA em *E. coli* tem a seguinte sequência de pares de nucleotídeos:

```

3'-ATGCTACTGCTATTCGCTGTATCG-5'
  |||
5'-TACGATGACGATAAGCGACATAGC-3'

```

Quando este segmento de DNA é transcrito pela RNA-polimerase, qual será a sequência de nucleotídeos no RNA transcrito se o promotor estiver situado à esquerda da sequência mostrada?

- 11.17.** Um segmento de DNA em *E. coli* tem a seguinte sequência de pares de nucleotídeos:

```

3'-ATATTACTGCAATGGGCTGTATCG-
  |||
5'-TATAATGACGTTACCCGACATAGC-
  |||
ATGCTACTGCTATTCGCTGTATCG-5'
  |||
TACGATGACGATAAGCGACATAGC-3'

```

Quando este segmento de DNA é transcrito pela RNA-polimerase, qual será a sequência de nucleotídeos no RNA transcrito?

- 11.18.** Um segmento de DNA em *E. coli* tem a seguinte sequência de pares de nucleotídeos:

```

3'-AACTGTACGTGCTACCTTGCTGATATTACT-
  |||
5'-TTGACATGCACGATGGAACGACTATAATGA-
  |||
GCAATGGGCTGTATCGATGCTACTGCTAT-5'
  |||
CGTTACCCGACATAGCTACGATGACGATA-3'

```

Quando este segmento de DNA é transcrito pela RNA-polimerase, qual será a sequência de nucleotídeos no RNA transcrito?

- 11.19.** Um segmento de DNA humano tem a seguinte sequência de pares de nucleotídeos:

```

3'-ATATTTACGTGCTACCTTGCTGATAGGACT-
|||||
5'-TATAAATGCACGATGGAACGACTATCCTGA-
GCAATGGGCTGTATCGATGCTACTGCTAT-5'
|||||
CGTTACCCGACATAGCTACGATGACGATA-3'

```

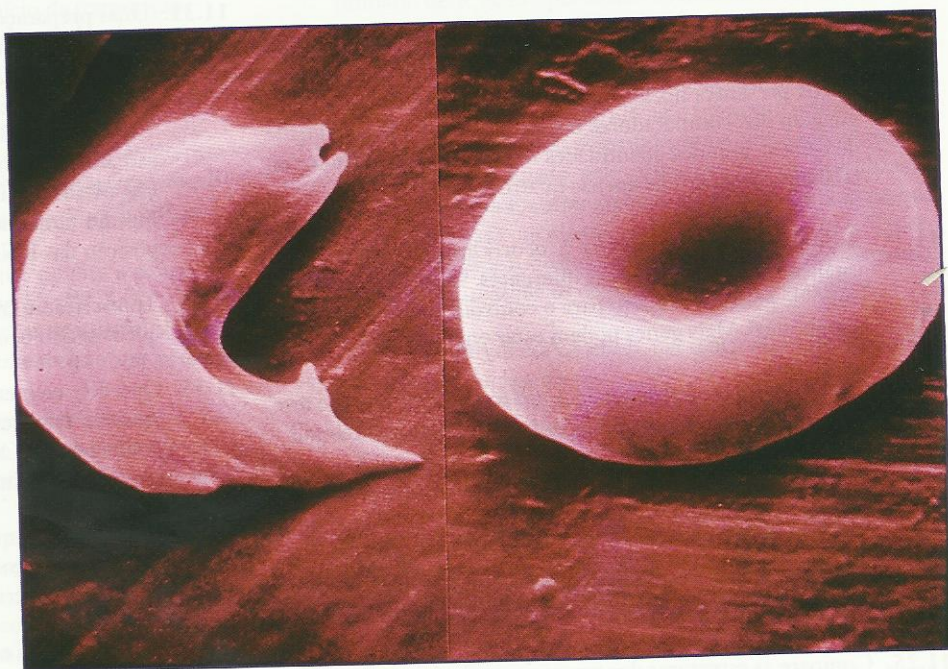
Quando este segmento de DNA é transcrito pela RNA-polimerase, qual será a sequência de nucleotídeos no RNA transcrito?

- 11.20.** O genoma de um humano deve estocar uma imensa quantidade de informações usando os quatro pares de nucleotídeos presentes no DNA. O que o código Morse e a linguagem dos computadores nos dizem sobre a possibilidade de estocar grandes quantidades de informações usando um alfabeto composto de apenas quatro letras?
- 11.21.** A biossíntese do metabólito X ocorre mediante seis etapas catalisadas por seis enzimas diferentes. Qual o número mínimo de genes necessário para o controle genético desta via metabólica? Mais genes podem estar envolvidos? Por quê?
- 11.22.** Qual o dogma central da genética molecular? Que impacto teve a descoberta dos vírus tumorais com RNA no dogma central?
- 11.23.** O que os processos de síntese de DNA, síntese de RNA e síntese de polipeptídeos têm em comum?
- 11.24.** Quais são os dois estágios da expressão gênica? Onde eles ocorrem em uma célula eucariótica? em uma célula procariótica?
- 11.25.** Compare as estruturas dos transcritos primários com as dos mRNA em procariontes e eucariontes. Em média, em que grupos de organismos eles diferem mais?
- 11.26.** Que cinco tipos de moléculas de RNA participam do processo de expressão gênica? Quais são as funções de cada tipo de RNA? Que tipos de RNA desempenham suas funções no (a) núcleo e (b) no citoplasma?
- 11.27.** Dois genes eucarióticos codificam dois polipeptídeos diferentes, cada um dos quais com 335 aminoácidos de tamanho. Um gene contém um único éxon; o outro gene contém um íntron com 41.324 pares de nucleotídeos de tamanho. Que gene você esperaria que fosse transcrito no menor tempo? Por quê? Quando os mRNA especificados por estes genes são traduzidos, que mRNA você esperaria que fosse traduzido no menor tempo? Por quê?
- 11.28.** Por que a necessidade de um RNA intermediário na síntese de proteínas é mais óbvia em eucariontes? Como os pesquisadores primeiro demonstraram que a síntese de RNA ocorreu no núcleo e a síntese de proteína ocorreu no citoplasma?
- 11.29.** Crie um experimento para demonstrar que os RNA transcritos são sintetizados no núcleo de eucariontes e são subsequentemente transportados para o citoplasma.
- 11.30.** RNA total foi isolado de células humanas crescidas em cultura. Este RNA foi misturado com filamentos não-moldes (unifilamentares) do gene humano que codifica a enzima timidina-cinase, e a mistura RNA-DNA foi incubada por 12 horas em condições de renaturação. Você esperaria que algum dúplice RNA-DNA fosse formado durante a incubação? Caso sim, por quê? Caso não, por que não? O mesmo experimento foi então feito usando-se o filamento molde do gene de timidina-cinase. Você esperaria que algum dúplice RNA-DNA fosse formado neste segundo experimento? Caso sim, por quê? Caso não, por que não?
- 11.31.** Duas preparações de RNA-polimerase de *E. coli* foram usadas em experimentos separados para catalisar a síntese de RNA *in vitro* usando-se um fragmento purificado de DNA que leva o gene *argH* (Fig. 11.1) como molde de DNA. Uma preparação catalisa a síntese de cadeias de RNA que são altamente heterogêneas em tamanho. A outra preparação catalisa a síntese de cadeias de RNA que são todas do mesmo tamanho. Qual a diferença mais provável na composição das RNA-polimerases nas duas preparações?
- 11.32.** Transcrição e tradução são acopladas em procariontes. Por que isto não ocorre em eucariontes?
- 11.33.** Que dois elementos estão quase sempre presentes nos promotores de genes eucarióticos que são transcritos pela RNA-polimerase II? Onde estes elementos estão situados com relação ao sítio de início da transcrição? Quais são suas funções?
- 11.34.** De que modos a maioria dos transcritos de genes eucarióticos são modificados? Quais são as funções destas modificações pós-transcricionais?
- 11.35.** De que modo a edição do RNA contribui para a diversidade de proteínas em eucariontes? Que papéis os RNA guias têm na edição do RNA?
- 11.36.** Como diferem os mecanismos de remoção dos íntrons de precursores de tRNA, precursores do rRNA de *Tetrahymena* e pré-mRNA nucleares? Em quais processos estão envolvidos os snRNA? Que papéis têm os snRNA?
- 11.37.** Que papel o gene da doença humana lúpus eritematoso sistêmico desempenham na caracterização dos snRNP humanos?
- 11.38.** Uma mutação em um gene humano essencial muda o sítio de corte em 5' de um grande íntron de GT para CC. Preveja o fenótipo de uma pessoa homozigota para esta mutação.
- 11.39.** RNA total foi isolado de núcleos de células humanas que crescem em cultura. Este RNA foi misturado com um fragmento desnaturado e purificado de DNA que leva um grande íntron de um gene de manutenção (um gene expresso em quase todas as células), e a mistura RNA-DNA foi incubada por 12 horas em condições de renaturação. Você esperaria que fosse formado algum dúplice RNA-DNA durante a incubação? Caso sim, por quê? Caso não, por que não? O mesmo experimento foi então feito usando-se o RNA citoplasmático total destas células. Você esperaria que fosse formado algum dúplice RNA-DNA neste segundo experimento? Caso sim, por quê? Caso não, por que não?

Tradução e o Código Genético

Tópicos

- Estrutura Protéica
- Síntese de Proteínas: Tradução
- O Código Genético
- Interações Códon-tRNA
- Confirmação *In Vivo* da Natureza do Código Genético



Micrografia eletrônica de varredura de hemácias falcêmica (à esquerda) e normal.

Anemia Falciforme: Efeitos Devastadores de uma Única Mudança de Par de Bases

Em 1904, James Herrick, um médico de Chicago, e Ernest Irons, um interno que trabalhava sob a supervisão de Herrick, examinaram as células sangüíneas de um de seus pacientes. Eles notaram que muitas das hemácias do jovem eram finas e alongadas, em marcante contraste com as hemácias redondas de seus outros pacientes. Eles obtiveram amostras frescas de sangue e repetiram seus exames microscópicos várias vezes, tendo sempre observado o mesmo resultado. O sangue deste paciente sempre continha células no formato das foices que os fazendeiros usavam, naquele tempo, para colher grãos.

O paciente era um estudante de 20 anos de idade que tinha períodos de fraqueza e tontura. Em muitos aspectos, o paciente parecia normal, tanto física quanto mentalmente. Seu principal problema era fadiga. Entretanto, um exame físico mostrou um coração aumentado e linfonodos aumentados. Seu coração sempre pareceu funcionar demais, mesmo quando ele estava em repouso. Testes sangüíneos mostraram que o paciente era anêmico; o conteú-

do de hemoglobina de seu sangue era cerca da metade do nível normal. Hemoglobina é o complexo protéico que leva oxigênio dos pulmões para os outros tecidos. Herrick registrou os sintomas de seu paciente por seis anos antes de publicar suas observações em 1910. Em sua publicação, Herrick deu ênfase à natureza crônica da anemia e à presença das hemácias falciformes. Em 1916, aos 32 anos, o paciente morreu de anemia grave e dano renal.

James Herrick foi o primeiro a publicar uma descrição da anemia falciforme, a primeira doença humana herdada a ser compreendida em nível molecular. A hemoglobina contém quatro polipeptídeos – duas cadeias de α -globina e duas cadeias de β -globina – e um grupo heme contendo ferro. Em 1957, Vernon Ingram e colaboradores demonstraram que o sexto aminoácido da cadeia β da hemoglobina falciforme era valina, enquanto o ácido glutâmico estava presente nesta posição na hemoglobina humana adulta normal. Esta única mudança de aminoácido em uma única cadeia polipeptídica é responsável por todos os sintomas da anemia falciforme.

Como a informação genética de um organismo, estocada na sequência de pares de nucleotídeos no DNA, controla o fenótipo de um organismo? Como a mudança de um par de nucleotídeos em um gene, como a mutação que causa anemia falciforme, altera a estrutura de uma proteína, o emissário pelo qual o gene atua? No Cap. 11, discutimos a transferência de informação genética estocada nas sequências de pares de nucleotídeos no DNA para as sequências de nucleotídeos em moléculas de mRNA, que, em eucariontes, levam esta informação do núcleo para os sítios de síntese de proteínas no citoplasma. A transferência de informação do DNA para o RNA (transcrição) e o processamento do RNA ocorrem no núcleo. Neste capítulo, examinamos o processo pelo qual a informação genética estocada em sequências de nucleotídeos nos mRNA é usada para especificar as sequências de aminoácidos em produtos gênicos polipeptídicos. Este processo, a **tradução**, ocorre no citoplasma em “bancadas” complexas chamadas ribossomos e requer a participação de muitas macromoléculas.

ESTRUTURA PROTÉICA

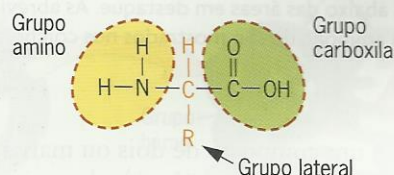
Proteínas são macromoléculas complexas compostas de 20 aminoácidos diferentes.

Coletivamente, as proteínas constituem cerca de 15% do peso total das células. As moléculas de água respondem por 70% do peso total das células vivas. Com a exceção da água, as proteínas são de longe o componente mais prevalente de organismos vivos em termos de massa total. Não só as proteínas são componentes importantes em termos de massa celular, mas também têm muitos papéis vitais para a vida de todas as células. Antes de discutir a síntese de proteínas, precisamos conhecer melhor sua estrutura.

Poliipeptídeos: Vinte Aminoácidos Diferentes

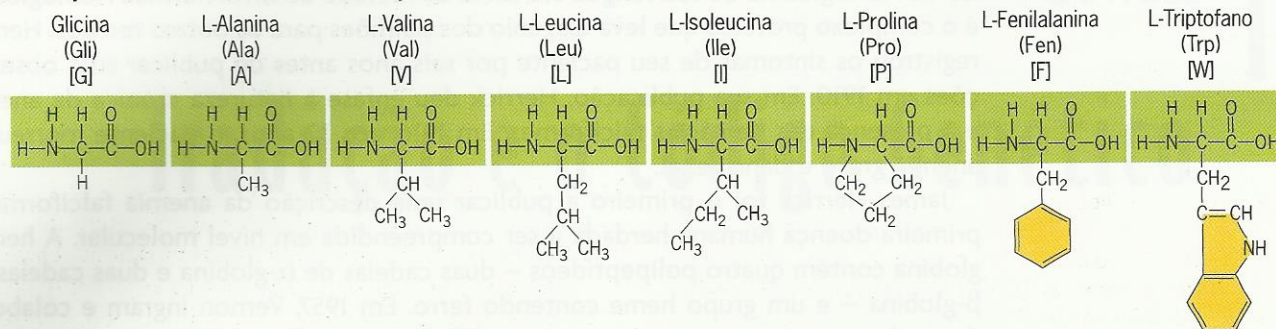
Como vimos no Cap. 11, as proteínas são compostas de polipeptídeos, e cada polipeptídeo é codificado por um gene. Cada polipeptídeo consiste em uma longa sequência de aminoácidos ligados por uniões covalentes. Vinte aminoácidos diferentes estão presentes na maioria das proteínas. Ocasio-

nalmente, um ou mais dos aminoácidos são quimicamente modificados após a síntese de um polipeptídeo, produzindo um novo aminoácido na proteína final. As estruturas dos 20 aminoácidos comuns são mostradas na Fig. 12.1. Todos os aminoácidos exceto a prolina contêm um **grupo amino livre** e um **grupo carboxila livre**.

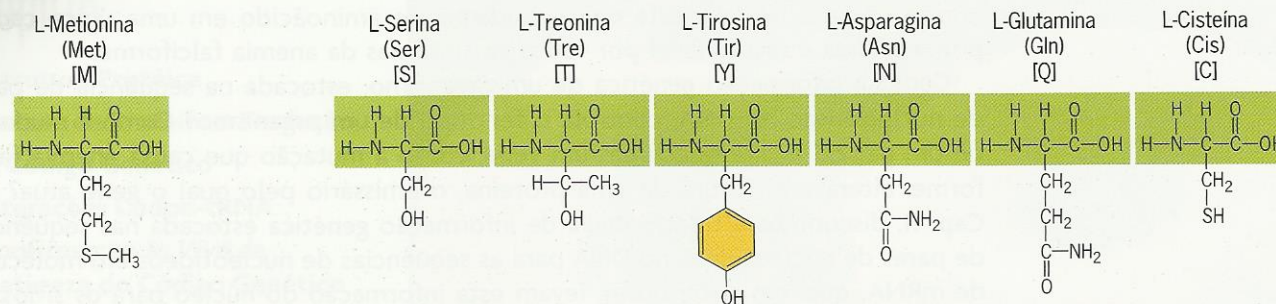


Os aminoácidos diferem uns dos outros pelos **grupos laterais** (chamados **R**, de radical) que estão presentes. Os grupos laterais altamente variados dão a diversidade estrutural das proteínas. Estas cadeias laterais são de quatro tipos: (1) grupos hidrofóbicos ou não polares, (2) grupos hidrofílicos ou polares, (3) grupos ácidos ou de carga negativa e (4) grupos básicos ou de carga positiva (Fig. 12.1). A diversidade química dos grupos laterais dos aminoácidos é responsável pela enorme versatilidade estrutural e funcional das proteínas.

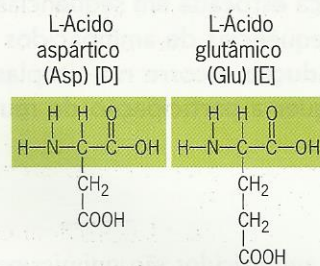
1. Grupos laterais hidrofóbicos ou não polares



2. Grupos laterais hidrofílicos ou polares



3. Grupos ácidos



4. Grupos básicos

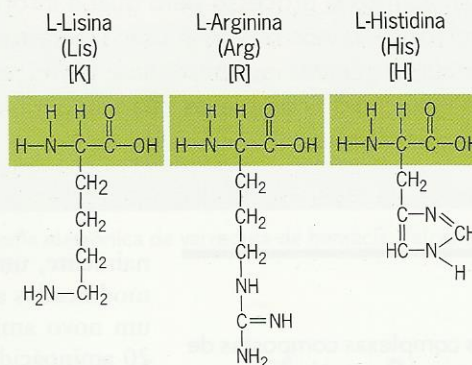


Fig. 12.1 ■ Estruturas dos 20 aminoácidos comumente encontrados em proteínas. Os grupos amino e carboxila, que participam de ligações peptídicas durante a síntese de proteínas, são mostrados nas áreas em destaque. Os grupos laterais, que são diferentes para cada aminoácido, são mostrados abaixo das áreas em destaque. As abreviações padrões com três letras são mostradas nos parênteses. Os símbolos de uma letra para cada aminoácido são mostrados nos colchetes.

Peptídeo é um composto de dois ou mais aminoácidos. Polipeptídeos são longas seqüências de aminoácidos, que variam de tamanho de 51 aminoácidos na insulina até mais de 1.000 aminoácidos na fibroína, a proteína da seda. Considerando os 20 aminoácidos diferentes comumente encontrados em polipeptídeos, o número de possíveis polipeptídeos diferentes é verdadeiramente imenso. Por exemplo, o número de seqüências diferentes de aminoácidos que podem ocorrer em um polipeptídeo contendo 100 aminoácidos é 20^{100} . Como 20^{100} é muito grande para compreender, consideremos um peptídeo pequeno. Existem 1,28 bilhão de seqüências diferentes de aminoácidos (20^7) possíveis em um

peptídeo com sete aminoácidos de tamanho. Os aminoácidos em polipeptídeos são ligados covalentemente por uniões chamadas **ligações peptídicas**. Cada ligação peptídica é formada por uma reação entre o grupo amino de um aminoácido e o grupo carboxila de um segundo aminoácido, com a eliminação de uma molécula de água (Fig. 12.2).

Proteínas: Estruturas Tridimensionais Complexas

Quatro níveis diferentes de organização – primário, secundário, terciário e quaternário – são distinguidos nas estruturas tridimensionais complexas das proteínas. A **estrutura**

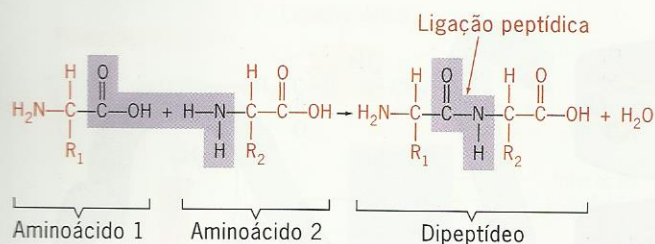


Fig. 12.2 ■ A formação de uma ligação peptídica entre dois aminoácidos pela remoção de água. Cada ligação peptídica conecta o grupo amino de um aminoácido com o grupo carboxila do aminoácido adjacente.

primária de um polipeptídeo é sua sequência de aminoácidos, que é especificada pela sequência de nucleotídeos de um gene. A **estrutura secundária** de um polipeptídeo refere-se às interrelações espaciais dos aminoácidos em segmentos do polipeptídeo. A **estrutura terciária** de um polipeptídeo refere-se a seu dobramento geral no espaço tridimensional e a **estrutura quaternária** refere-se à associação de dois ou mais polipeptídeos em uma proteína multimérica. A hemoglobina fornece um exemplo excelente da complexidade das proteínas, exibindo os quatro níveis de organização estrutural (Fig. 12.3).

A maioria dos polipeptídeos dobra-se espontaneamente em uma conformação específica ditada por suas estruturas primárias. Caso desnaturado (desdobrado) por tratamento com solventes apropriados, a maioria das proteínas reconstituirá suas conformações originais quando é removido o agente desnaturante. Assim, na maioria dos casos, toda a informação necessária para a determinação da forma reside na estrutura primária da proteína. Em alguns casos, o dobramento da proteína envolve interações com proteínas cha-

mas **chaperonas** que ajudam os polipeptídeos nascentes a formar a estrutura tridimensional apropriada.

Os dois tipos mais comuns de estruturas secundárias em proteínas são as **α -hélices** e as **folhas β** (Fig. 12.4). Ambas as estruturas são mantidas por pontes de hidrogênio entre ligações peptídicas situadas bem próximas umas das outras. A α -hélice é um cilindro rígido no qual cada ligação peptídica faz pontes de hidrogênio com a ligação peptídica entre aminoácidos distantes três a quatro aminoácidos. Devido a sua estrutura rígida, a prolina não pode estar presente dentro de uma α -hélice (Fig. 12.4a). Uma folha β ocorre quando um polipeptídeo se dobra sobre si mesmo, às vezes repetidamente, e os segmentos paralelos são mantidos no lugar por pontes de hidrogênio entre ligações peptídicas vizinhas (Fig. 12.4b).

Embora a organização espacial de aminoácidos adjacentes e segmentos de um polipeptídeo determine sua estrutura secundária, o dobramento geral do polipeptídeo completo define sua estrutura terciária, ou **conformação**. Em geral, aminoácidos com cadeias laterais hidrofílicas estão situados nas superfícies de proteínas (em contato com o citoplasma aquoso), enquanto aqueles com cadeias laterais hidrofóbicas interagem entre si nas regiões interiores. A estrutura terciária de uma proteína é mantida primariamente por um grande número de ligações não covalentes relativamente fracas. As únicas ligações covalentes que têm um papel significativo na conformação da proteína são as pontes dissulfídicas (S—S) que se formam entre cisteínas apropriadamente posicionadas (Fig. 12.5). Entretanto, estão envolvidos quatro tipos diferentes de interações não covalentes: (1) ligações iônicas, (2) pontes de hidrogênio, (3) interações hidrofóbicas e (4) interações de van der Waals (Fig. 12.5).

Ligações iônicas ocorrem entre cadeias laterais de aminoácidos de cargas opostas – por exemplo, os grupos laterais de

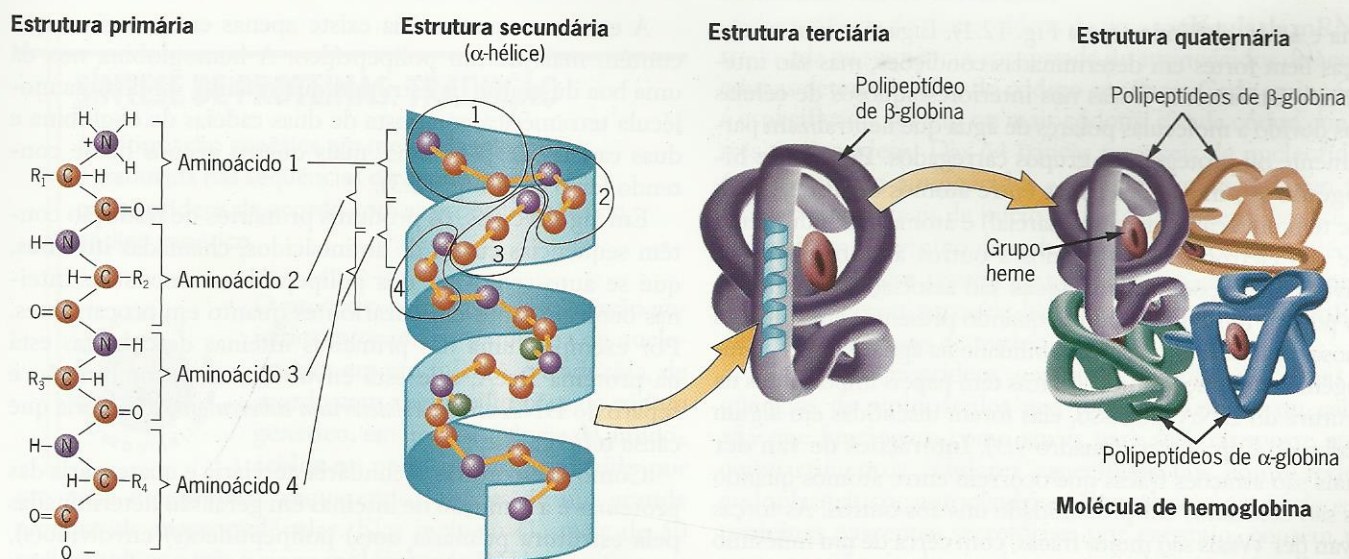


Fig. 12.3 ■ Os quatro níveis de organização em proteínas – (1) primária, (2) secundária, (3) terciária e (4) quaternária – são ilustrados usando-se a hemoglobina humana como um exemplo.

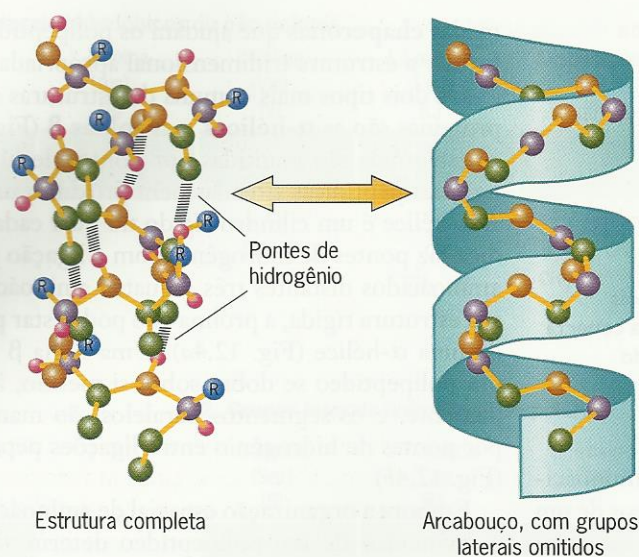
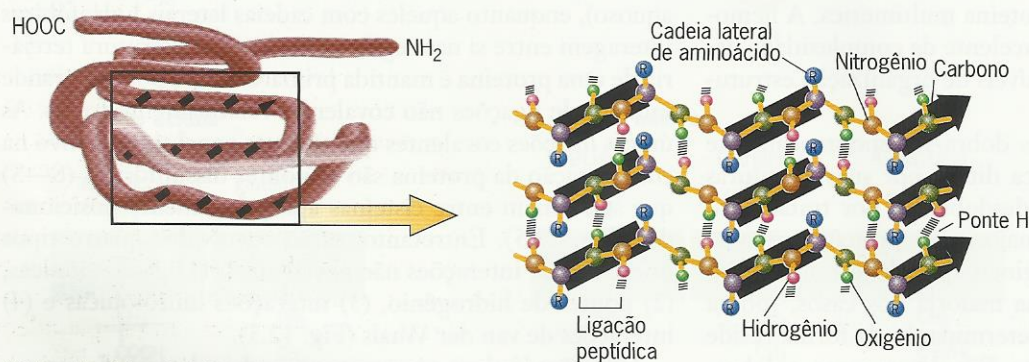
(a) Estrutura de uma região de α -hélice de um polipeptídeo.(b) Região típica de folha β de um polipeptídeo.

Fig. 12.4 ■ Estrutura secundária em proteínas. A α -hélice mostrada é da mioglobina, a proteína transportadora de oxigênio no músculo de vertebrados; a folha β ilustrada aqui é de uma molécula de imunoglobulina (anticorpo) humana.

lisina e ácido glutâmico (veja Fig. 12.1). Ligações iônicas são forças bem fortes em determinadas condições, mas são interações relativamente fracas nos interiores aquosos de células vivas devido a moléculas polares de água que neutralizam parcialmente ou protegem os grupos carregados. **Pontes de hidrogênio** são interações fracas entre átomos eletronegativos (que têm uma carga negativa parcial) e átomos de hidrogênio (que são eletropositivos) ligados a outros átomos eletronegativos. **Interações hidrofóbicas** são associações de grupos não polares uns com os outros quando presentes em soluções aquosas em razão de sua insolubilidade na água. Pontes de hidrogênio e interações hidrofóbicas têm papéis importantes na estrutura do DNA; por isso, elas foram discutidas em algum detalhe no Cap. 9 (veja Quadro 9.3). **Interações de van der Waals** são atrações fracas que ocorrem entre átomos quando eles são colocados em proximidade uns aos outros. As forças de van der Waals são muito fracas, com cerca de um milésimo da força de uma ligação covalente, mas têm um papel importante em manter as conformações de regiões proximamente alinhadas de macromoléculas.

A estrutura quaternária existe apenas em proteínas que contêm mais de um polipeptídeo. A hemoglobina nos dá uma boa ilustração da estrutura quaternária, sendo uma molécula tetramérica composta de duas cadeias de α -globina e duas cadeias de β -globina, mais quatro grupos heme contendo ferro (veja Fig. 12.3).

Em alguns casos, os produtos primários de tradução contêm seqüências curtas de aminoácidos, chamadas **inteínas**, que se auto-removem dos polipeptídeos nascentes. Inteínas ocorrem tanto em eucariontes quanto em procariontes. Por exemplo, uma das primeiras inteínas descobertas está na proteína RecA, que está envolvida em recombinação e reparo do DNA em *Mycobacterium tuberculosis*, a bactéria que causa tuberculose.

Como as estruturas secundária, terciária e quaternária das proteínas e a remoção de inteínas em geral são determinadas pela estrutura primária do(s) polipeptídeo(s) envolvido(s), no restante deste capítulo enfocaremos os mecanismos pelos quais os genes controlam a estrutura primária dos polipeptídeos.

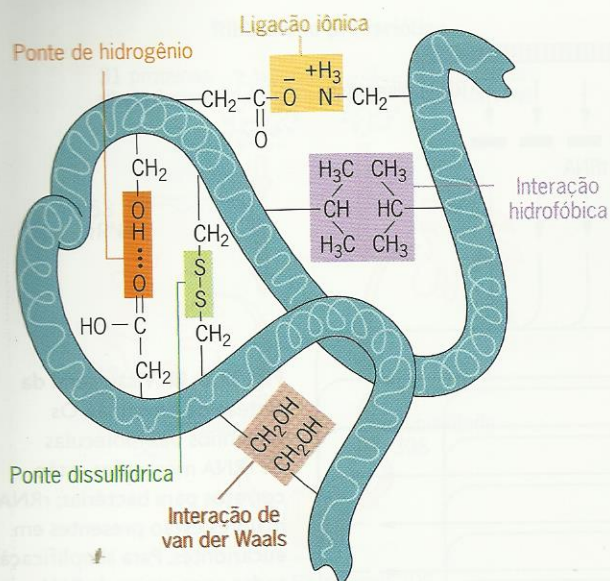


Fig. 12.5 ■ Os cinco tipos de interações moleculares que determinam a estrutura terciária, ou conformação tridimensional, de um polipeptídeo. A ponte dissulfídica é uma ligação covalente; todas as outras interações são não-covalentes.

PONTOS IMPORTANTES

- A maioria dos genes exercem seus efeitos no fenótipo de um organismo por meio de proteínas, que são grandes macromoléculas compostas de polipeptídeos.
- Cada polipeptídeo é um polímero do tipo cadeia composto de aminoácidos diferentes.
- A sequência de aminoácidos de cada polipeptídeo é especificada pela sequência de nucleotídeos de um gene.
- A grande diversidade funcional de proteínas resulta em parte de suas complexas estruturas tridimensionais.

SÍNTESE DE PROTEÍNAS: TRADUÇÃO

A informação genética em moléculas de mRNA é traduzida nas sequências de aminoácidos de polipeptídeos de acordo com as especificações do código genético.



O processo pelo qual a informação genética estocada na sequência de nucleotídeos em um mRNA é traduzida, de acordo com as especificações do código genético, em uma sequência de aminoácidos no polipeptídeo produzido por

um gene é complexo, requerendo funções de um grande número de macromoléculas. Elas incluem (1) mais de 50 polipeptídeos e três a cinco moléculas de RNA presentes em cada ribossomo (a composição exata varia entre espécies), (2) pelo menos 20 enzimas ativadoras de aminoácidos, (3)

40 a 60 moléculas diferentes de tRNA e (4) várias proteínas solúveis envolvidas em iniciação, alongamento e término da cadeia polipeptídica. Como muitas destas macromoléculas, particularmente os componentes do ribossomo, estão presentes em grandes quantidades em cada célula, o sistema de tradução constitui parte importante da maquinaria metabólica de cada célula.

Visão Geral da Síntese de Proteínas

Antes de focar os detalhes do processo de tradução, devemos rever o processo de síntese de proteínas em sua totalidade. Uma visão geral da síntese de proteínas, ilustrando sua complexidade e as principais macromoléculas envolvidas, é apresentada na Fig. 12.6. A primeira etapa na expressão gênica, a transcrição, envolve a transferência da informação estocada nos genes para intermediários de RNA mensageiro (mRNA), que levam a informação para os sítios de síntese de polipeptídeos no citoplasma. A transcrição é discutida em detalhe no Cap. 11. A segunda etapa, a tradução, envolve a transferência da informação nas moléculas de mRNA para as sequências de aminoácidos nos produtos gênicos polipeptídicos.

A tradução ocorre nos ribossomos, que são estruturas macromoleculares complexas situadas no citoplasma. A tradução envolve três tipos de RNA, todos os quais são transcritos a partir de moldes de DNA (genes cromossômicos). Além dos mRNA, três a cinco moléculas de RNA (moléculas de rRNA) estão presentes como parte da estrutura de cada ribossomo, e 40 a 60 moléculas de pequeno RNA (moléculas de tRNA) funcionam como adaptadores mediando a incorporação dos aminoácidos apropriados em polipeptídeos em resposta a sequências específicas de nucleotídeos nos mRNA. Os aminoácidos são ligados a moléculas corretas de tRNA por um conjunto de enzimas ativadoras chamadas **aminoacil-tRNA-sintetases**.

A sequência de nucleotídeos de uma molécula de mRNA é traduzida na sequência apropriada de aminoácidos de acordo com as determinações do código genético. Cada aminoácido é especificado por um ou mais códons, e cada códon contém três nucleotídeos. Das 64 trincas possíveis de nucleotídeos, 61 especificam aminoácidos – duas das quais também funcionam como sinais de início de cadeia polipeptídica – e três especificam o término da cadeia polipeptídica. As moléculas de tRNA contêm trincas de nucleotídeos chamadas anticódons, que fazem pares de bases com os códons no mRNA durante o processo de tradução.

Alguns polipeptídeos nascentes contêm pequenas sequências de aminoácidos nos terminais amino ou carboxila que funcionam como sinais para seu transporte até os compartimentos celulares específicos tais como retículo endoplasmático, mitocôndrias, cloroplastos ou núcleos. As proteínas nascentes secretórias, por exemplo, contêm uma curta **sequência de sinal** no terminal amino que direciona o polipeptídeo emergente para as membranas do retículo endoplasmático. Sequências alvos similares estão presentes

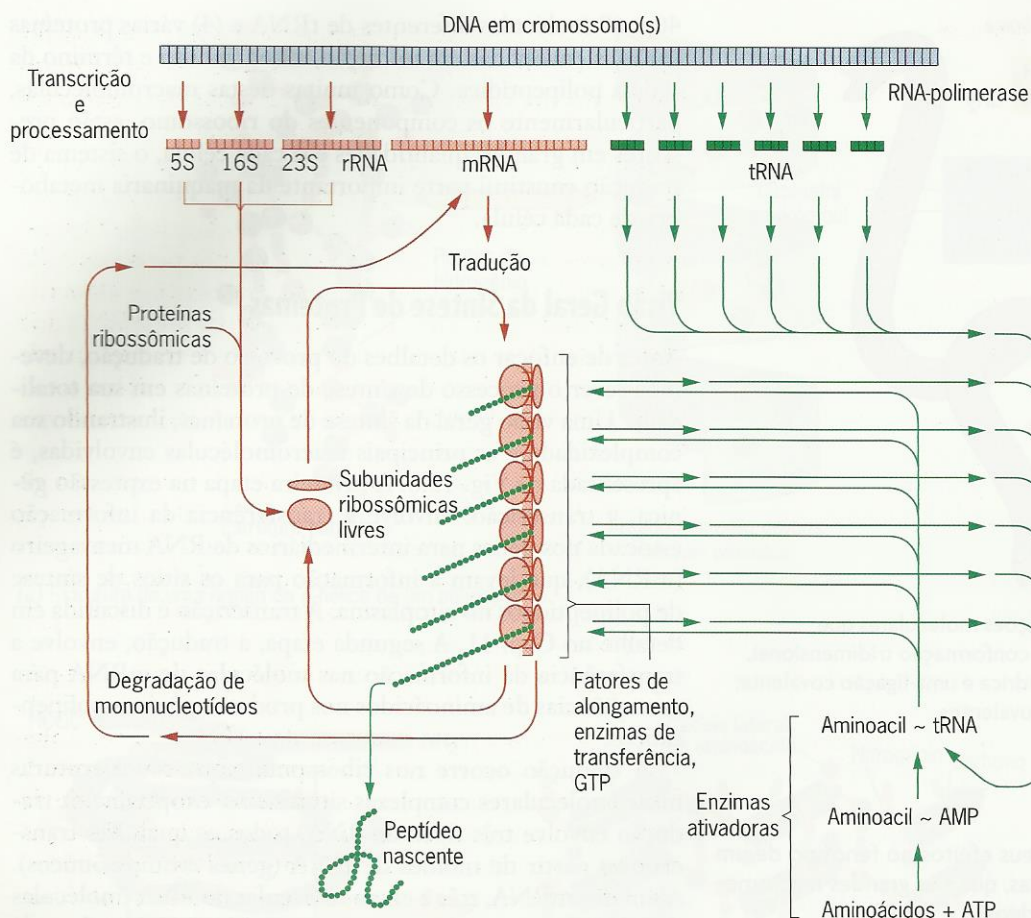


Fig. 12.6 ■ Visão geral da síntese de proteínas. Os tamanhos das moléculas de rRNA mostrados estão corretos para bactérias; rRNA maiores estão presentes em eucariontes. Para simplificação, todas as espécies de RNA foram transcritas de segmentos contíguos de uma única molécula de DNA. Na verdade, os vários RNA são transcritos de genes situados em partes diferentes de um a muitos cromossomos. Detalhes dos vários estágios da síntese de proteínas são discutidos em partes subsequentes deste capítulo.

nos terminais amino das proteínas destinadas a mitocôndrias e cloroplastos. Algumas proteínas nucleares contêm extensões direcionais nas pontas carboxila. Em muitos casos, os peptídeos direcionais são removidos enzimaticamente por peptidases específicas após o transporte da proteína para o compartimento celular apropriado.

Os ribossomos podem ser vistos como “bancadas”, completas com máquinas e ferramentas necessárias para fazer um polipeptídeo. Eles são inespecíficos no sentido de que podem sintetizar qualquer polipeptídeo (qualquer sequência de aminoácidos) codificado por uma determinada molécula de mRNA, mesmo um mRNA de uma espécie diferente. Cada molécula de mRNA é traduzida simultaneamente por vários ribossomos, resultando na formação de um polirribossomo, ou polissomo. Feita esta breve revisão da síntese de proteínas, examinaremos alguns dos componentes mais importantes da maquinaria de tradução mais atentamente.

Componentes Necessários para a Síntese de Proteínas: Ribossomos e RNA Transportadores

As células vivas devotam mais energia à síntese de proteínas do que a qualquer outro aspecto do metabolismo. Cerca de um terço da massa seca total da maioria das células consiste

em moléculas que participam diretamente da biossíntese de proteínas. Em *E. coli*, os aproximadamente 200.000 ribossomos correspondem a 25% da massa seca de cada célula. Esse envolvimento de uma grande parte da maquinaria metabólica das células com o processo de síntese de proteínas documenta sua importância nas formas de vida que existem em nosso planeta.

Quando os sítios de síntese de proteínas foram marcados em células cultivadas por curtos intervalos na presença de aminoácidos radioativos e foram visualizados por auto-radiografia, os resultados mostraram que as proteínas são sintetizadas nos ribossomos. Em procariontes, os ribossomos são distribuídos pelas células. Em eucariontes, eles estão situados no citoplasma, freqüentemente na ampla rede intracelular de membranas do retículo endoplasmático.

Os ribossomos são aproximadamente metade proteína e metade RNA (Fig. 12.7). Eles são compostos de duas subunidades, uma maior e uma menor, que se dissociam quando a tradução de uma molécula de mRNA está completa, e se reassociam durante o início da tradução. Os tamanhos dos ribossomos são mais freqüentemente expressos em termos de suas velocidades de sedimentação durante centrifugação, em unidades Svedberg (veja Enfoque Técnico: Técnicas de Centrifugação, no Cap. 10). O ribossomo de *E. coli*, como os ribossomos de outros procariontes, tem peso molecular

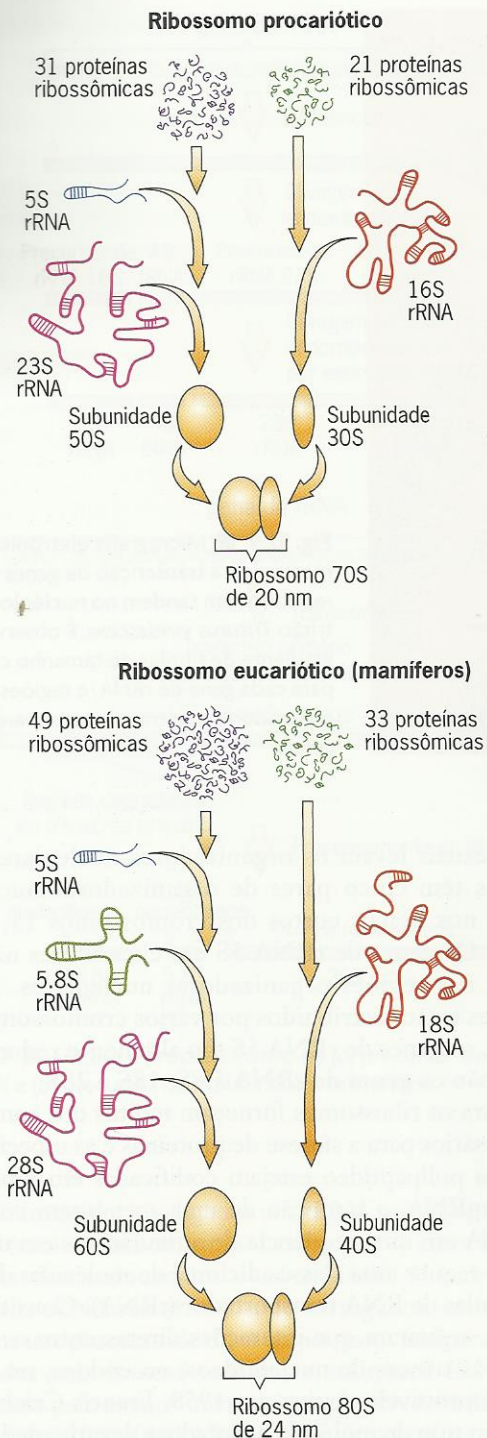


Fig. 12.7 ■ Composição macromolecular de ribossomos procarióticos e eucarióticos.

de $2,5 \times 10^6$, tamanho de 70S e dimensões de cerca de 20 nm \times 25 nm. Os ribossomos de eucariontes são maiores (em geral cerca de 80S); entretanto, o tamanho varia entre espécies. Os ribossomos presentes nas mitocôndrias e cloroplastos de células eucarióticas são menores (geralmente cerca de 60S).

Embora o tamanho e a composição macromolecular dos ribossomos variem, a estrutura tridimensional geral do ribossomo é basicamente a mesma em todos os organismos. Em *E. coli*, a subunidade ribossômica menor (30S) contém uma molécula de RNA 16S (peso molecular de cerca de 6×10^5) mais 21 polipeptídeos diferentes, e a subunidade maior (50S) contém duas moléculas de RNA (5S, peso molecular de cerca de 4×10^4 , e 23S, peso molecular de cerca de $1,2 \times 10^6$) mais 31 polipeptídeos. Em ribossomos de mamíferos, a subunidade menor contém uma molécula de RNA 18S mais 33 polipeptídeos, e a subunidade maior contém três moléculas de RNA com os tamanhos 5S, 5,8S e 28S mais 49 polipeptídeos. Em organelas, os tamanhos correspondentes de rRNA são 5S, 13S e 21S.

Masayasu Nomura e colaboradores foram capazes de desmontar a subunidade 30S de *E. coli* em macromoléculas individuais e então reconstituir subunidades 30S funcionais a partir dos componentes. Deste modo, eles estudaram as funções de rRNA individuais e moléculas de proteínas ribossômicas.

As moléculas de RNA ribossômicas, como as moléculas de mRNA, são transcritas a partir de um molde de DNA. Em eucariontes, a síntese de rRNA ocorre no nucléolo (Fig. 12.8) e é catalisada pela RNA-polimerase I. O nucléolo é um componente altamente especializado do núcleo dedicado exclusivamente à síntese de rRNA e sua montagem em ribossomos. Os genes de RNA ribossômicos estão presentes em arranjos em tandem duplicados separados por regiões espaçadoras intergênicas. A transcrição destes arranjos em tandem de genes para rRNA pode ser visualizada diretamente por microscopia eletrônica (Fig. 12.9).

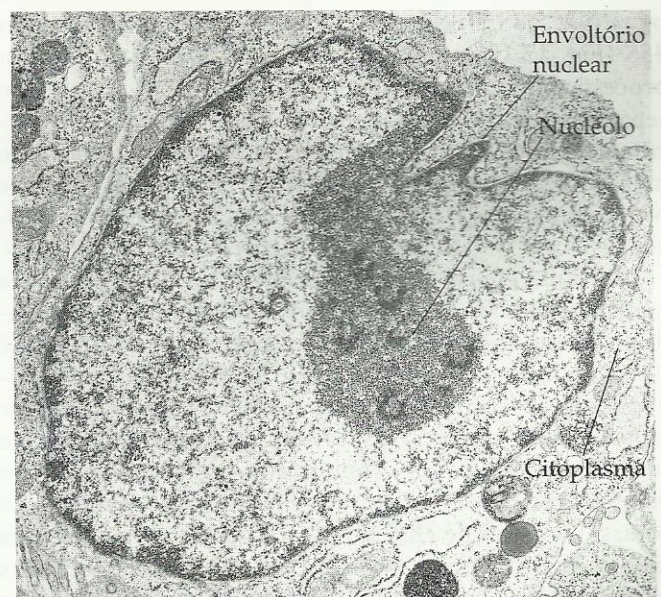


Fig. 12.8 ■ Micrografia eletrônica de um fibroblasto humano, mostrando o nucléolo, o local de síntese de rRNA e montagem do ribossomo.

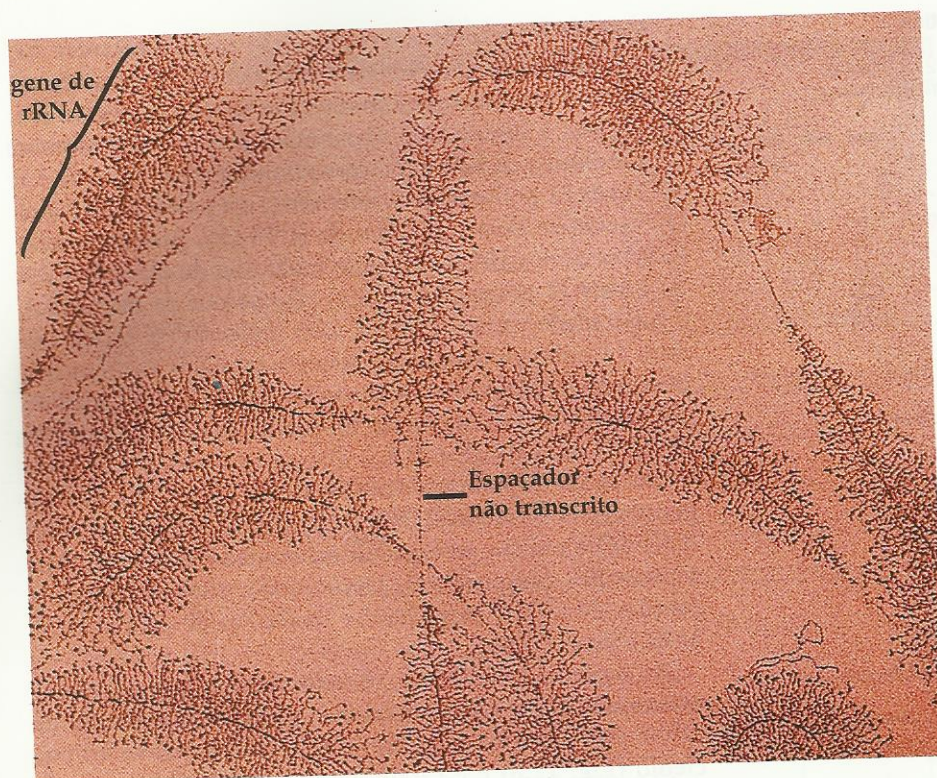


Fig. 12.9 ■ Micrografia eletrônica mostrando a transcrição de genes de rRNA repetidos em tandem no nucléolo do tritão *Triturus viridescens*. É observado um gradiente de fibrilas de tamanho crescente para cada gene de rRNA, e regiões espaçadoras não transcritas separam os genes.

A transcrição dos genes de rRNA produz precursores que são muito maiores que as moléculas de RNA encontradas em ribossomos. Esses precursores de rRNA sofrem processamento pós-transcricional para produzir as moléculas finais de rRNA. Em *E. coli*, o transcrito gênico de rRNA é um precursor 30S, que sofre clivagens endonucleolíticas para produzir os rRNA 5S, 16S e 23S mais uma molécula de RNA transportador 4S (Fig. 12.10a). Em mamíferos, os rRNA 5,8S, 18S e 28S são clivados de um precursor 45S (Fig. 12.10b), enquanto o rRNA 5S é produzido por processamento pós-transcricional de um transcrito gênico separado. Além das clivagens pós-transcricionais dos precursores de rRNA, muitos dos nucleotídeos nos rRNA são metilados pós-transcricionalmente. A metilação parece proteger moléculas de rRNA de degradação por ribonucleases.

Múltiplas cópias dos genes para rRNA estão presentes nos genomas de todos os organismos que foram estudados até hoje. Esta redundância de genes de rRNA não é surpreendente considerando-se o grande número de ribossomos presentes por célula. Em *E. coli*, sete genes de rRNA (*rrnA*—*rrnE*, *rrnG*, *rrnH*) são distribuídos entre três sítios distintos no cromossomo. Em eucariontes, os genes de rRNA estão presentes em centenas a milhares de cópias. Os genes de rRNA 5,8S, 18S e 28S de eucariontes estão presentes em disposições em tandem nas **regiões organizadoras nucleolares** de cromossomos. Em alguns eucariontes, tais como milho, há um único par de organizadores nucleolares (no cromossomo 6 em milho). Em *Drosophila* e no sapo sul-africano *Xenopus laevis*, os cromos-

somos sexuais levam os organizadores nucleolares. Seres humanos têm cinco pares de organizadores nucleolares situados nos braços curtos dos cromossomos 13, 14, 15, 21 e 22. Os genes de rRNA 5S em eucariontes não estão situados nas regiões organizadoras nucleolares. Em vez disso, eles estão distribuídos por vários cromossomos. Entretanto, os genes de rRNA 5S são altamente redundantes, como o são os genes de rRNA 5,8S, 18S e 28S.

Embora os ribossomos forneçam muitos dos componentes necessários para a síntese de proteínas e as especificações para cada polipeptídeo estejam codificadas em uma molécula de mRNA, a tradução de uma mensagem codificada no mRNA em uma sequência de aminoácidos em um polipeptídeo requer uma classe adicional de moléculas de RNA, as moléculas de RNA transportador (tRNA). Considerações químicas sugeriram que interações diretas entre os aminoácidos e as trincas de nucleotídeos, ou códons, no mRNA seriam improváveis. Assim, em 1958, Francis Crick propôs que algum tipo de molécula adaptadora deveria mediar a especificação de aminoácidos por códons no mRNA durante a síntese de proteínas. As moléculas adaptadoras foram logo identificadas por outros pesquisadores e demonstraram ser pequenas moléculas de RNA (4S, com 70 a 95 nucleotídeos de tamanho). Essas moléculas, primeiro chamadas moléculas de RNA solúvel (sRNA) e subsequentemente moléculas de RNA transportador (tRNA), contêm uma sequência tripla de nucleotídeos, o anticódon, que é complementar e faz pares de bases com a sequência códon no mRNA durante a tradução. Existem de um a quatro tRNA para cada um dos 20 aminoácidos.

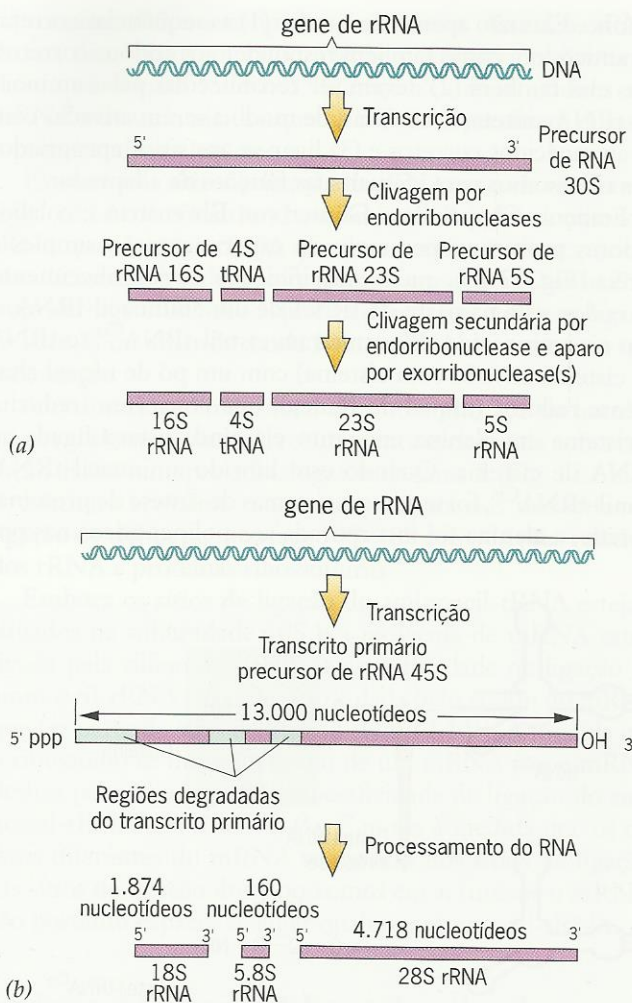
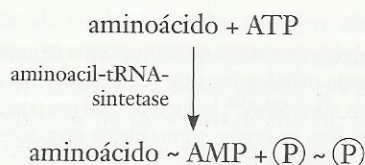


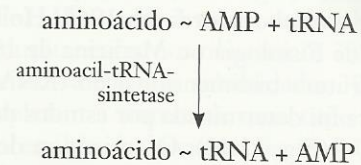
Fig. 12.10 ■ Síntese e processamento de (a) precursor de rRNA 30S em *E. coli* e (b) precursor de rRNA 45S em mamíferos.

Os aminoácidos são ligados aos tRNA por ligações de alta energia (muito reativas, simbolizadas por ~) entre os grupos carboxila dos aminoácidos e os terminais 3'-hidroxila dos tRNA. Os tRNA são ativados, ou carregados, com aminoácidos em um processo de duas etapas, com ambas as reações catalisadas pela mesma enzima, a aminoacil-tRNA-sintetase. Existe pelo menos uma aminoacil-tRNA-sintetase para cada um dos 20 aminoácidos. A primeira etapa na síntese de aminoacil-tRNA envolve a ativação do aminoácido usando energia do trifosfato de adenosina (ATP):



O aminoácido ~ AMP intermediário não é normalmente liberado da enzima antes de sofrer a segunda etapa na síntese

de aminoacil-tRNA, ou seja, a reação com o tRNA apropriado:



Os aminoacil ~ tRNA são substratos para a síntese de polipeptídeos em ribossomos, com cada tRNA ativado reconhecendo o códon correto de mRNA e apresentando o aminoácido em uma configuração estérica (estrutura tridimensional) que facilita a formação de ligações peptídicas.

Os tRNA são transcritos de genes. Como no caso dos rRNA, os tRNA são transcritos sob a forma de moléculas precursoras maiores que sofrem processamento pós-transcricional (clivagem, aparo, metilação e assim em diante). As moléculas finais de tRNA contêm vários nucleosídeos que não estão presentes nos transcritos primários de genes de tRNA. Esses nucleosídeos incomuns, tais como inosina, pseudo-uridina, diidrouridina, 1-metilguanossina e vários outros, são produzidos por modificações pós-transcricionais catalisadas por enzimas dos quatro nucleosídeos incorporados ao RNA durante a transcrição.

Devido a seu pequeno tamanho (a maioria tem de 70 a 95 nucleotídeos de tamanho), os tRNA são mais propensos a análise estrutural que outras moléculas maiores de RNA

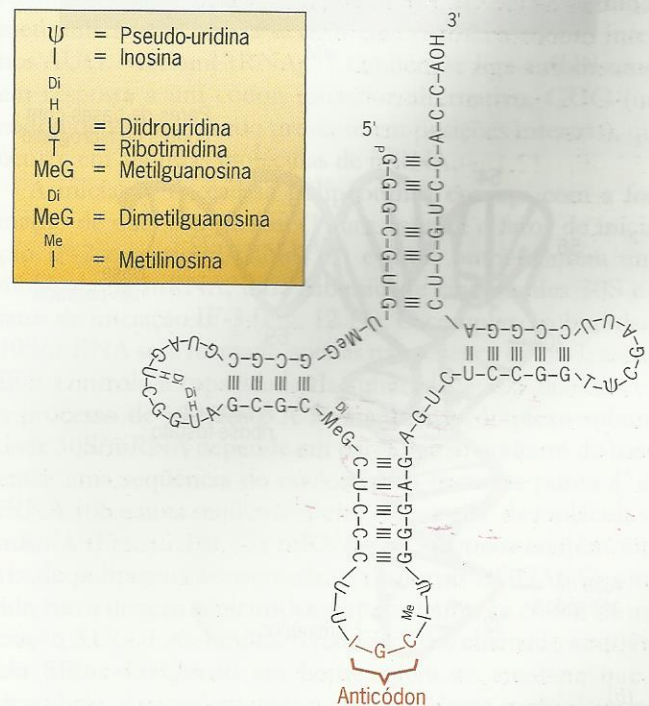
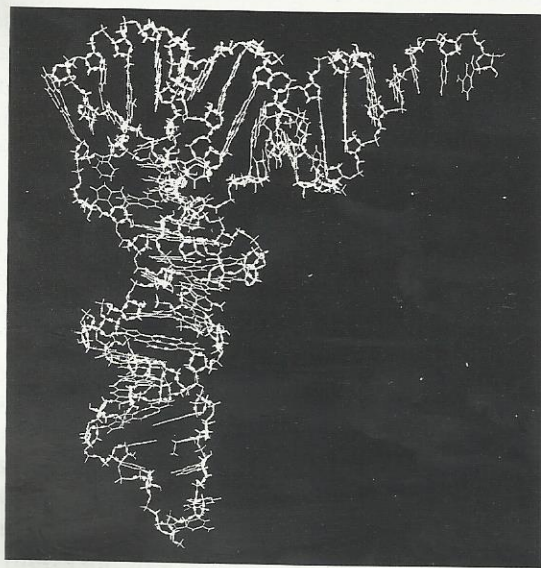


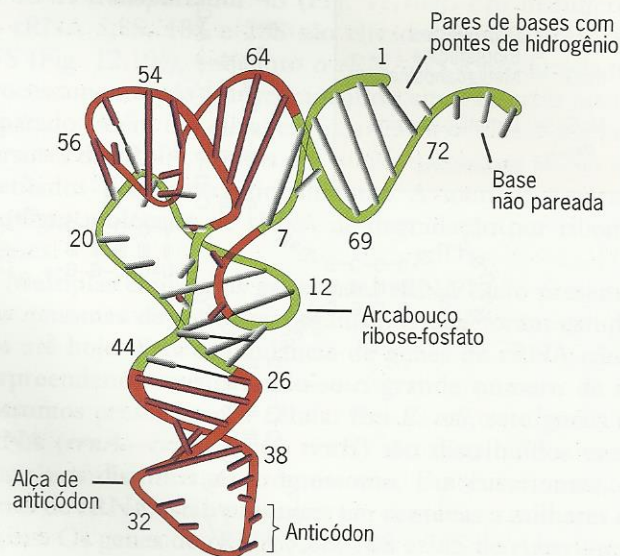
Fig. 12.11 ■ Sequência de nucleotídeos e configuração em trevo do tRNA de alanina de *S. cerevisiae*. Os nomes dos nucleotídeos modificados presentes no tRNA são mostrados no detalhe.

envolvidas na síntese de proteínas. A sequência completa de nucleotídeos e a proposta estrutura em trevo do tRNA de alanina de levedura (Fig. 12.11) foram publicadas por Robert W. Holley e colaboradores em 1965. Holley dividiu o Prêmio Nobel de Fisiologia ou Medicina de 1968 pelo seu trabalho. A estrutura tridimensional do tRNA de fenilalanina de levedura foi determinada por estudos de difração de raios X em 1974 (Fig. 12.12). O anticódon de cada tRNA ocorre dentro de uma alça (região sem pontes de hidrogênio) perto do meio da molécula.

Deve ser aparente que as moléculas de tRNA devem conter uma grande especificidade a despeito de seu pequeno tamanho.



(a)

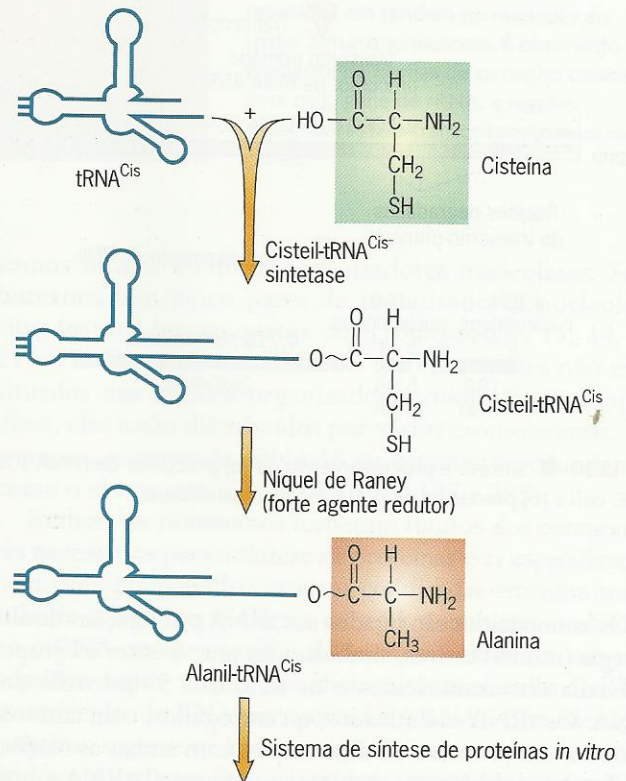


(b)

Fig. 12.12 ■ Foto (a) e desenho interpretativo (b) de um modelo molecular do tRNA de fenilalanina de levedura baseado em dados de difração de raios X.

manho. Elas não apenas devem ter (1) as sequências corretas de anticódon, como também responder aos códons corretos, mas elas também (2) devem ser reconhecidas pelas aminoacil-tRNA-sintetases corretas, de modo a serem ativadas com os aminoácidos corretos e (3) ligar-se aos sítios apropriados nos ribossomos para efetuar suas funções de adaptador.

François Chapeville e Günter von Ehrenstein e colaboradores provaram, por meio de experimentação simples e direta (Fig. 12.13), que a especificidade de reconhecimento de códon está na parte do tRNA de um aminoacil-tRNA, e não no aminoácido. Eles trataram cisteil-tRNA^{Cis} (o tRNA de cisteína ativado com cisteína) com um pó de níquel altamente redutor (níquel de Raney), que converteu (reduziu) a cisteína em alanina enquanto ela ainda estava ligada ao tRNA de cisteína. Quando este híbrido aminoacil-tRNA, alanil-tRNA^{Cis}, foi usado em sistemas de síntese de proteínas *in vitro*, a alanina foi incorporada aos polipeptídeos nas po-



Experimento I:

Usando poli(UG) (UGUGU etc., co-polímero UG repetido) como um RNA mensageiro artificial. A alanina ligada ao tRNA^{Cis} foi incorporada, a despeito do fato de que os códons de alanina são GCU, GCC, GCA e GCG. UGU = códon de cisteína.

Experimento II:

Usando o sistema de síntese de hemoglobina de reticulócito de coelho contendo mRNA nativos de hemoglobina. Foi demonstrado que alanina de alanil-tRNA^{Cis} foi incorporada em posições nas cadeias de globina de coelho normalmente ocupadas por cisteína.

Fig. 12.13 ■ Prova de que a especificidade de reconhecimento de códon de um complexo aminoacil-tRNA reside em um tRNA, e não no aminoácido.

sições normalmente ocupadas por cisteína. Assim, os tRNA de fato funcionam como as moléculas adaptadoras que Crick disse que deviam mediar a interação entre os códons nos mRNA e os aminoácidos que os códons especificam durante o processo de tradução.

Existem três sítios de ligação de tRNA em cada ribossomo (Fig. 12.14a-b). O **sítio A** ou **aminoacil** liga o aminoacil-tRNA que chega, o tRNA que traz o aminoácido seguinte a ser adicionado à cadeia polipeptídica crescente. O **sítio P** ou **peptidil** liga o tRNA ao qual a cadeia crescente está ligada. O **sítio E** ou **sítio de saída** liga o tRNA descarregado que está saindo.

A estrutura tridimensional do ribossomo 70S da bactéria *Thermus thermophilus* foi recentemente esclarecida com resolução de 0,55 nm por cristalografia de raios X (Fig. 12.14c-e). A estrutura cristalina mostra as posições dos três sítios de ligação de tRNA na interface 50S-30S e as posições relativas dos rRNA e proteínas ribossômicas.

Embora os sítios de ligação do aminoacil-tRNA estejam situados na subunidade 50S e a molécula de mRNA esteja ligada pela subunidade 30S, a especificidade de ligação do aminoacil-tRNA em cada sítio é dada pelo códon do mRNA que faz parte do sítio de ligação (Fig. 12.14b). À medida que o ribossomo se move ao longo de um mRNA (ou o mRNA desliza pelo ribossomo), a especificidade de ligação do aminoacil-tRNA nos sítios A, P e E muda à medida que os códons diferentes do mRNA se movem nos sítios de ligação. Os sítios de ligação dos ribossomos em si (menos o mRNA) são portanto capazes de ligar qualquer aminoacil-tRNA.

Tradução: A Síntese de Polipeptídeos Usando Moldes de mRNA

Revisamos todos os principais componentes do sistema de síntese de proteínas. As moléculas de mRNA fornecem as especificações para as seqüências de aminoácidos dos polipeptídeos codificados pelo gene. Os ribossomos fornecem muitos dos componentes macromoleculares necessários para o processo de tradução. Os tRNA fornecem as moléculas adaptadoras necessárias para incorporar aminoácidos em polipeptídeos em resposta a códons nos mRNA. Além disso, várias proteínas solúveis participam do processo. A tradução da seqüência de nucleotídeos em uma molécula de mRNA na seqüência de aminoácidos de seu produto polipeptídico pode ser dividida em três etapas: (1) iniciação da cadeia polipeptídica, (2) alongamento da cadeia e (3) término da cadeia.

A **iniciação** da tradução inclui todos os eventos que precedem a formação de uma ligação peptídica entre os dois primeiros aminoácidos da nova cadeia polipeptídica. Embora vários aspectos do processo de iniciação sejam os mesmos em procariontes e eucariontes, alguns são diferentes. Examinaremos primeiro a iniciação das cadeias polipeptídicas em *E. coli* e veremos então os aspectos únicos do início da tradução em eucariontes.

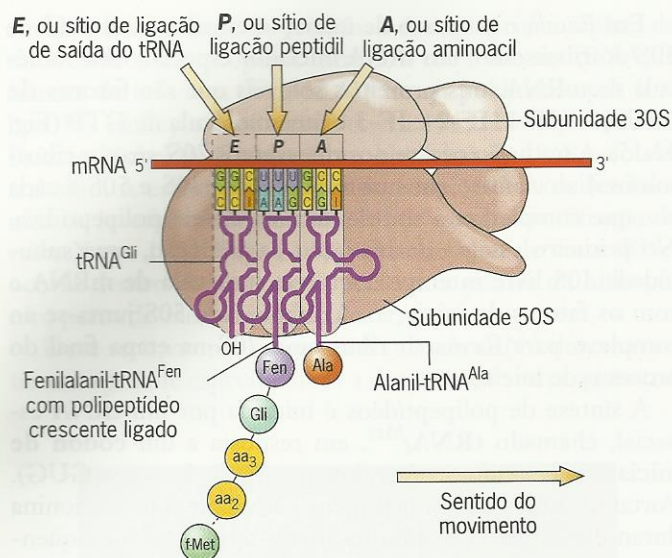
Em *E. coli*, o processo de iniciação envolve a subunidade 30S do ribossomo, um tRNA iniciador especial, uma molécula de mRNA, três proteínas solúveis que são **fatores de iniciação: IF-1, IF-2 e IF-3** e uma molécula de GTP (Fig. 12.15). A tradução ocorre nos ribossomos 70S, mas os ribossomos dissociam-se em suas subunidades 30S e 50S a cada vez que completam a síntese de uma cadeia polipeptídica. No primeiro estágio da iniciação da tradução, uma subunidade 30S livre interage com uma molécula de mRNA e com os fatores de iniciação. A subunidade 50S junta-se ao complexo para formar o ribossomo 70S na etapa final do processo de iniciação.

A síntese de polipeptídeos é iniciada por um tRNA especial, chamado $\text{tRNA}_{\text{f}}^{\text{Met}}$, em resposta a um **códon de iniciação** da tradução (geralmente AUG, às vezes GUG). Portanto, todos os polipeptídeos começam com metionina durante a síntese. A metionina amino-terminal é subsequentemente removida de muitos polipeptídeos. Assim, proteínas funcionais não precisam ter uma metionina amino-terminal. A metionina do $\text{tRNA}_{\text{f}}^{\text{Met}}$ tem o grupo amino bloqueado

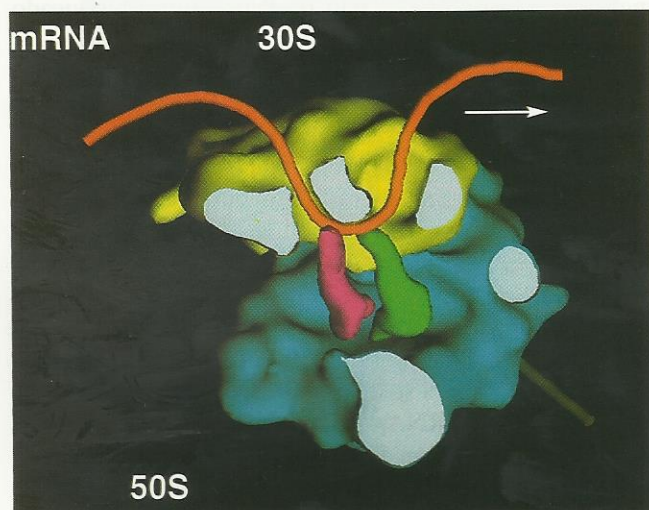


por um grupo formil ($-\text{C}-\text{H}$) (donde o subscrito "f" no $\text{tRNA}_{\text{f}}^{\text{Met}}$). Uma metionina tRNA diferente, tRNA_{Met} , responde a códons internos de metionina. Ambos os tRNA de metionina têm o mesmo anticódon e ambos respondem ao mesmo códon (AUG) para metionina. Entretanto, apenas metionil- $\text{tRNA}_{\text{f}}^{\text{Met}}$ interage com o fator IF-2 de iniciação da proteína para começar o processo de iniciação (Fig. 12.15). Assim, apenas metionil- $\text{tRNA}_{\text{f}}^{\text{Met}}$ liga-se ao ribossomo em resposta a códons de iniciação AUG nos mRNA, deixando metionil- tRNA_{Met} para se ligar em resposta a códons internos AUG. Metionil- $\text{tRNA}_{\text{f}}^{\text{Met}}$ também se liga a ribossomos em resposta a um códon iniciador alternativo, GUG (um códon de valina quando presente em posições internas), que ocorre em algumas moléculas de mRNA.

A iniciação da cadeia polipeptídica começa com a formação de dois complexos: (1) um contém o fator de iniciação IF-2 e metionil- $\text{tRNA}_{\text{f}}^{\text{Met}}$, e (2) o outro contém uma molécula de mRNA, uma subunidade ribossômica 30S e o fator de iniciação IF-3 (Fig. 12.15). O complexo subunidade 30S/mRNA será formado apenas na presença de IF-3; assim, IF-3 controla a capacidade da subunidade 30S de começar o processo de iniciação. A formação do complexo subunidade 30S/mRNA depende em parte do pareamento de bases entre uma seqüência de nucleotídeos perto da ponta 3' do rRNA 16S e uma seqüência perto da ponta 5' da molécula de mRNA (Fig. 12.16). Os mRNA procarióticos contêm uma via de polipurina conservada, o consenso AGGAGG, situado cerca de sete nucleotídeos antecedentes ao códon de iniciação AUG. Este hexâmero conservado, chamado **seqüência Shine-Dalgarno** em homenagem ao cientista que o descobriu, é complementar a uma seqüência perto da ponta 3' do RNA ribossômico 16S. Quando as seqüências Shine-Dalgarno dos mRNA são experimentalmente modificadas de modo a não poder fazer mais pares de bases com o rRNA



(a) Diagrama do ribossomo 70S



(b) Ribossomo 70S – modelo visto em corte



(c) Ribossomo 70S – estrutura cristalina



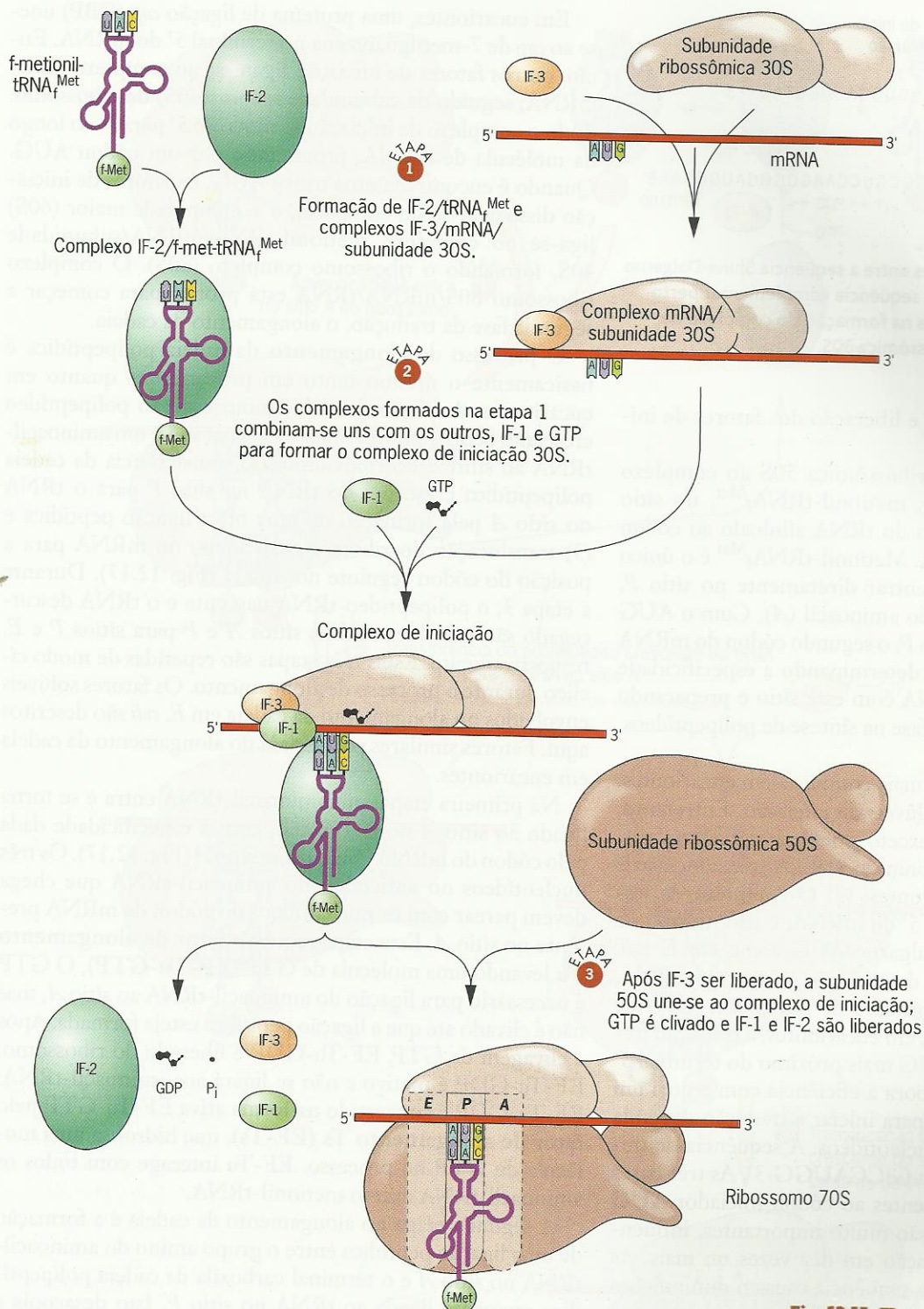
(d) Subunidade 50S – estrutura cristalina



(e) Subunidade 30S – estrutura cristalina

Fig. 12.14 ■ Estrutura de um ribossomo em *E. coli* (a–b) e *Thermus thermophilus* (c–e). (a) Cada complexo ribossomo/mRNA contém três sítios de ligação aminoacil-tRNA. O sítio A ou aminoacil-tRNA é ocupado por alanil-tRNA^{Ala}. O sítio P ou peptidil é ocupado por fenilalanil-tRNA^{Fen}, com a cadeia polipeptídica crescente ligada covalentemente ao tRNA de fenilalanina. O sítio E ou de saída é ocupado pelo tRNA^{Gli} antes de sua liberação do ribossomo. (b) Uma molécula de mRNA (laranja), que é ligada à subunidade 30S (verde-claro) do ribossomo, contribui para a especificidade dos sítios de ligação do tRNA, que estão situados amplamente na subunidade 50S (azul) do ribossomo. Os aminoacil-tRNA situados nos sítios P e A são mostrados em vermelho e verde-escuro, respectivamente. O sítio E está desocupado.

(c–e) Estrutura cristalina do ribossomo 70S com 0,55 nm de resolução, mostrando o ribossomo completo (c) e as interfaces das subunidades 50S (d) e 30S (e). (c) A subunidade 50S à esquerda; a subunidade 30S à direita. (d–e) Interfaces da subunidade 50S e 30S obtidas girando-se as estruturas mostradas em (c) 90° para a esquerda (d) ou para a direita (e), respectivamente. Os tRNA nos sítios A, P e E são mostrados em dourado, laranja e vermelho, respectivamente. Componentes: rRNA 16S (ciano); rRNA 23S (cinza); rRNA 5S (azul-claro); proteínas da subunidade 30S (azul-escuro); e proteínas da subunidade 50S (magenta). L1, grande subunidade de proteína 1; S7, pequena subunidade de proteína 7.

Fig. 12.15 ■ A iniciação da tradução em *E. coli*.

16S, os mRNA modificados ou não são traduzidos ou são traduzidos muito ineficientemente, indicando que esse pareamento de bases tem um papel importante na tradução.

O complexo IF-2/metionil-tRNA^{Met}_f e o complexo mRNA/subunidade 30S/IF-3 se combinam subsequentemente um com o outro e com o fator de iniciação IF-1 e uma molécula de GTP para formar o complexo de inicia-

ção 30S completo. A etapa final na iniciação da tradução é a adição da subunidade 50S ao complexo de iniciação 30S para produzir o ribossomo 70S completo. O fator de iniciação IF-3 deve ser liberado do complexo antes que a subunidade 50S possa juntar-se ao complexo. Nunca se observou associação de IF-3 e da subunidade 50S com a subunidade 30S ao mesmo tempo. A adição da subunidade

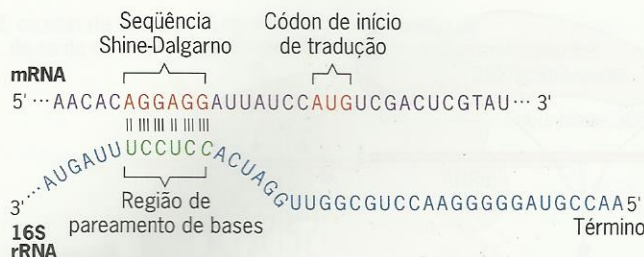


Fig. 12.16 ■ Pareamento de bases entre a sequência Shine-Dalgarno em um mRNA procariótico e uma sequência complementar perto da ponta 3' do rRNA 16S envolvida na formação do complexo de iniciação mRNA/subunidade ribossômica 30S.

50S requer energia do GTP e liberação dos fatores de iniciação IF-1 e IF-2.

A adição da subunidade ribossômica 50S ao complexo posiciona o tRNA iniciador, metionil-tRNA^{Met}, no sítio peptidil (*P*) com o anticódon do tRNA alinhado ao códon de iniciação AUG do mRNA. Metionil-tRNA^{Met} é o único aminoacil-tRNA que pode entrar diretamente no sítio *P*, sem primeiro passar pelo sítio aminoacil (*A*). Com o AUG iniciador posicionado no sítio *P*, o segundo códon do mRNA está posicionado no sítio *A*, determinando a especificidade de ligação do aminoacil-tRNA com este sítio e preparando as condições para a segunda fase na síntese de polipeptídeos, o alongamento da cadeia.

A iniciação da tradução é mais complexa em eucariontes, envolvendo vários fatores solúveis de iniciação. Entretanto, o processo geral é similar exceto por duas características. (1) O grupo amino da metionina no tRNA iniciador não é formilado como em procariontes. (2) O complexo de iniciação forma-se no terminal 5' do mRNA, e não no sítio de início de tradução Shine-Dalgarno/AUG como em *E. coli*. Em eucariontes, o complexo de iniciação percorre o mRNA, começando na ponta 5', à procura de um códon AUG de iniciação da tradução. Assim, em eucariontes, a tradução frequentemente começa no AUG mais próximo do terminal 5' da molécula de mRNA, embora a eficiência com a qual um determinado AUG é usado para iniciar a tradução dependa da sequência contígua de nucleotídeos. A sequência de iniciação ótima é 5'-GCC(A ou G)CCAUGG-3'. As três bases purínicas (A ou G) antecedentes ao códon iniciador AUG e o G imediatamente após são muito importantes, influenciando a eficiência de iniciação em dez vezes ou mais. As mudanças de outras bases na sequência causam diminuições menores na eficiência da iniciação. Os requisitos de sequência para uma iniciação ótima da tradução em eucariontes são denominados **regras de Kozak**, em homenagem a Marilyn Kozak, que primeiro as propôs.

Como os procariontes, os eucariontes contêm um tRNA iniciador especial, o tRNA^{Met} ("i" de iniciador), mas o grupo amino do metionil-tRNA^{Met} não é formilado. O metionil-tRNA^{Met} iniciador interage com um fator solúvel de iniciação e entra no sítio *P* diretamente durante o processo de iniciação, como em *E. coli*.

Em eucariontes, uma proteína de ligação *cap* (CBP) une-se ao *cap* de 7-metilguanosina no terminal 5' do mRNA. Então, outros fatores de iniciação ligam-se ao complexo CBP-mRNA, seguido da subunidade menor (40S) do ribossomo. Todo o complexo de iniciação se move de 5' para 3' ao longo da molécula de mRNA, procurando por um códon AUG. Quando é encontrada uma trinca AUG, os fatores de iniciação dissociam-se do complexo, e a subunidade maior (60S) liga-se ao complexo metionil-tRNA/mRNA/subunidade 40S, formando o ribossomo completo (80S). O complexo ribossomo 80S/mRNA/tRNA está pronto para começar a segunda fase da tradução, o alongamento da cadeia.

O processo de **alongamento** da cadeia polipeptídica é basicamente o mesmo tanto em procariontes quanto em eucariontes. A adição de cada aminoácido ao polipeptídeo crescente ocorre em três etapas: (1) ligação de um aminoacil-tRNA ao sítio *A* do ribossomo, (2) transferência da cadeia polipeptídica crescente do tRNA no sítio *P* para o tRNA no sítio *A* pela formação de uma nova ligação peptídica e (3) translocação do ribossomo ao longo do mRNA para a posição do códon seguinte no sítio *A* (Fig. 12.17). Durante a etapa 3, o polipeptídeo-tRNA nascente e o tRNA descarregado são translocados dos sítios *A* e *P* para sítios *P* e *E*, respectivamente. Estas três etapas são repetidas de modo cíclico durante o processo de alongamento. Os fatores solúveis envolvidos no alongamento da cadeia em *E. coli* são descritos aqui. Fatores similares participam no alongamento da cadeia em eucariontes.

Na primeira etapa, um aminoacil-tRNA entra e se torna ligado ao sítio *A* no ribossomo, com a especificidade dada pelo códon do mRNA alinhado ao sítio *A* (Fig. 12.17). Os três nucleotídeos no anticódon do aminoacil-tRNA que chega devem parear com os nucleotídeos do códon de mRNA presente no sítio *A*. Esta etapa requer o **fator de alongamento Tu** levando uma molécula de GTP (**EF-Tu-GTP**). O GTP é necessário para ligação do aminoacil-tRNA ao sítio *A*, mas não é clivado até que a ligação peptídica esteja formada. Após a clivagem de GTP, EF-Tu-GDP é liberado do ribossomo. EF-Tu-GDP é inativo e não se ligará aos aminoacil-tRNA. EF-Tu-GDP é convertido na forma ativa EF-Tu-GTP pelo **fator de alongamento Ts** (**EF-Ts**), que hidrolisa uma molécula de GTP no processo. EF-Tu interage com todos os aminoacil-tRNA exceto metionil-tRNA.

A segunda etapa no alongamento da cadeia é a formação de uma ligação peptídica entre o grupo amino do aminoacil-tRNA no sítio *A* e o terminal carboxila da cadeia polipeptídica crescente ligada ao tRNA no sítio *P*. Isto desacopla a cadeia crescente do tRNA no sítio *P* e liga covalentemente a cadeia ao tRNA no sítio *A* (Fig. 12.17). Esta reação importante é catalisada pela **peptidiltransferase**, uma atividade enzimática da subunidade 50s do ribossomo. Devemos notar que a atividade da peptidiltransferase reside na molécula de rRNA 23S e não em uma proteína ribossômica, talvez uma outra relíquia de um mundo primordial baseado em RNA. A formação da ligação peptídica requer a hidrólise da molécula de GTP levada ao ribossomo por EF-Tu na etapa 1.

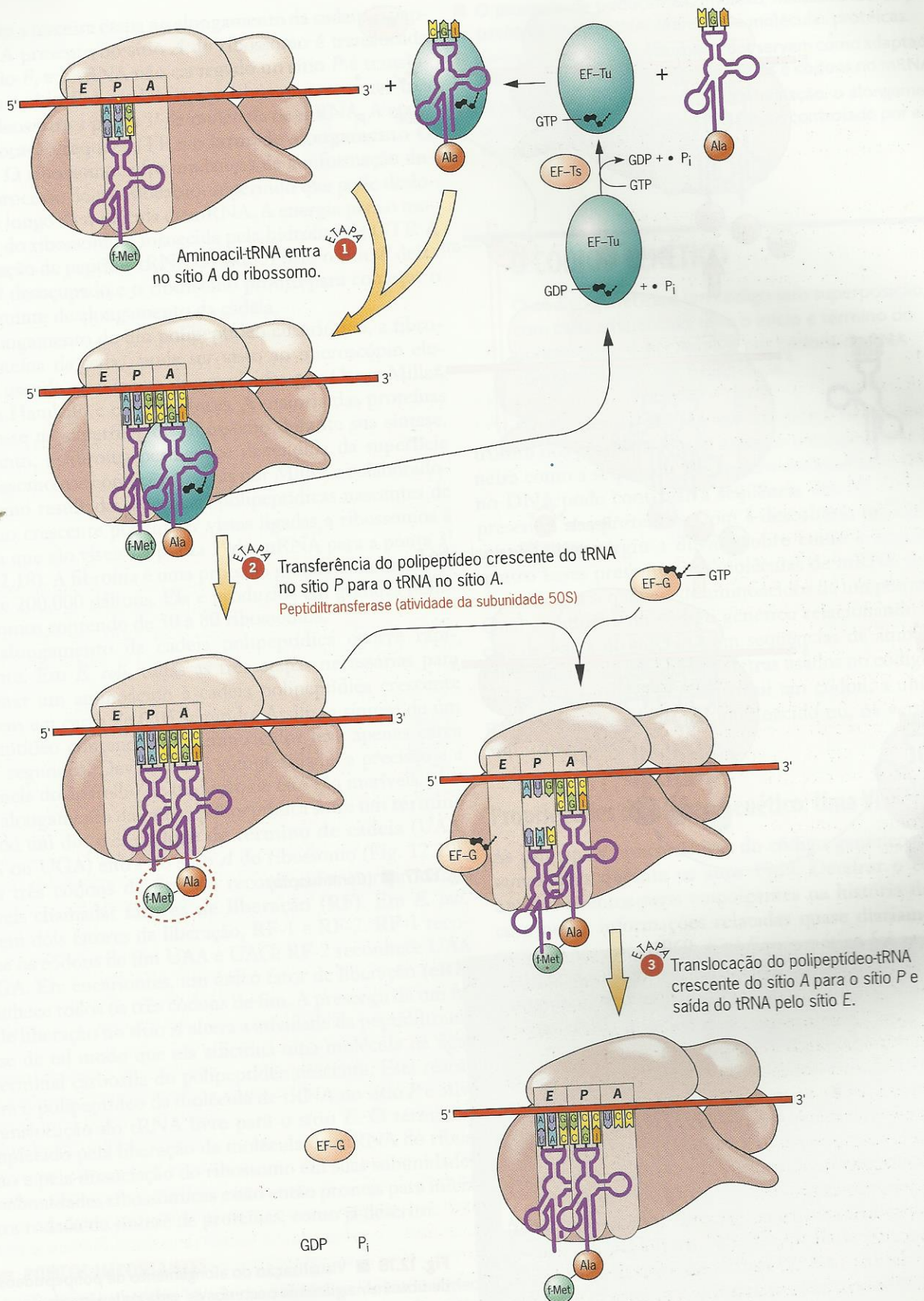


Fig. 12.17 ■ Alongamento da cadeia polipeptídica em *E. coli*.

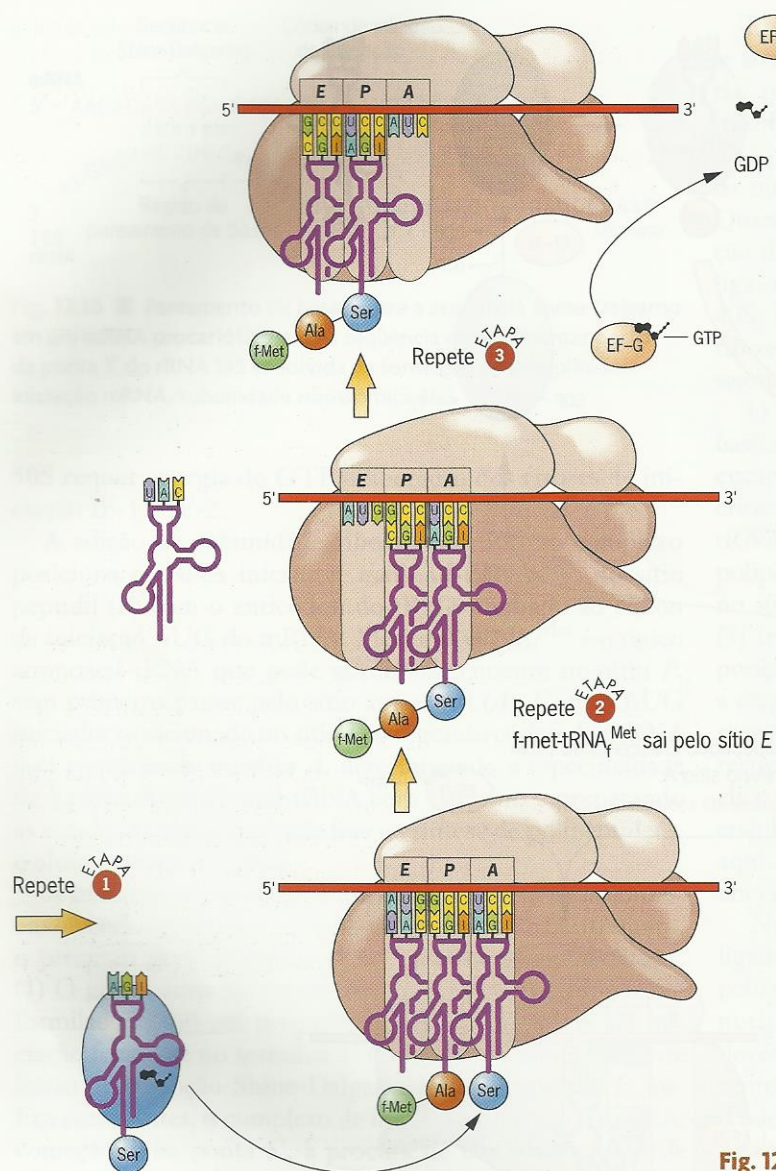
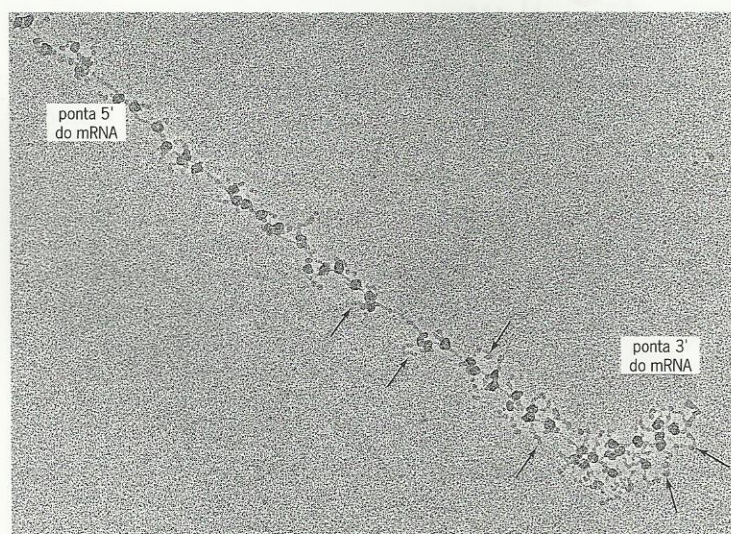


Fig. 12.17 ■ (Continuação)

Fig. 12.18 ■ Visualização do alongamento de polipeptídeos de fibroína na glândula posterior de seda da lagarta de *Bombyx mori*. As setas indicam crescimento dos polipeptídeos de fibroína. Note seu aumento de tamanho como um dos enfoques da ponta 3' da molécula de mRNA.

Durante a terceira etapa no alongamento da cadeia, o peptidil-tRNA presente no sítio *A* do ribossomo é translocado para o sítio *P*, e o tRNA não carregado no sítio *P* é translocado para o sítio *E*, à medida que o ribossomo se move três nucleotídeos para a ponta 3' da molécula de mRNA. A etapa de translocação requer GTP e o **fator de alongamento G (EF-G)**. O ribossomo sofre mudanças de conformação durante o processo de translocação, sugerindo que pode deslocar-se ao longo da molécula de mRNA. A energia para o movimento do ribossomo é fornecida pela hidrólise de GTP. A translocação de peptidil-tRNA do sítio *A* para o sítio *P* deixa o sítio *A* desocupado e o ribossomo pronto para começar o ciclo seguinte de alongamento da cadeia.

O alongamento de um polipeptídeo eucariótico, a fibroína (proteína da seda), pode ser visto ao microscópio eletrônico usando-se técnicas desenvolvidas por Oscar Miller, Barbara Hamkalo e colaboradores. A maioria das proteínas dobram-se na superfície do ribossomo durante sua síntese. Entretanto, a fibroína permanece distendida da superfície do ribossomo nas condições usadas por Miller e colaboradores. Como resultado, as cadeias polipeptídicas nascentes de tamanho crescente podem ser vistas ligadas a ribossomos à medida que são vistas da ponta 5' do mRNA para a ponta 3' (Fig. 12.18). A fibroína é uma proteína grande com massa de mais de 200.000 daltos. Ela é produzida em grandes polirribossomos contendo de 50 a 80 ribossomos.

O alongamento da cadeia polipeptídica ocorre rapidamente. Em *E. coli*, todas as três etapas necessárias para adicionar um aminoácido à cadeia polipeptídica crescente ocorrem em cerca de 0,05 segundo. Assim, a síntese de um polipeptídeo contendo 300 aminoácidos leva apenas cerca de 15 segundos. Devido a sua complexidade, a precisão e a eficiência do aparelho de tradução de fato são incríveis.

O alongamento da cadeia polipeptídica sofre um **término** quando um dos três **códons de término de cadeia** (UAA, UAG ou UGA) entra no sítio *A* do ribossomo (Fig. 12.19). Estes três códons de fim são reconhecidos por proteínas solúveis chamadas **fatores de liberação (RF)**. Em *E. coli*, existem dois fatores de liberação, RF-1 e RF-2. RF-1 reconhece os códons de fim UAA e UAG; RF-2 reconhece UAA e UGA. Em eucariontes, um único fator de liberação (**eRF**) reconhece todos os três códons de fim. A presença de um fator de liberação no sítio *A* altera a atividade da peptidiltransferase de tal modo que ela adiciona uma molécula de água ao terminal carboxila do polipeptídeo nascente. Esta reação libera o polipeptídeo da molécula de tRNA no sítio *P* e ativa a translocação do tRNA livre para o sítio *E*. O término é completado pela liberação da molécula de mRNA do ribossomo e pela dissociação do ribossomo em suas subunidades. As subunidades ribossômicas estão então prontas para iniciar outra rodada de síntese de proteínas, como já descrito.

■ PONTOS IMPORTANTES

- A informação genética levada nas seqüências de nucleotídeos nas moléculas de mRNA é traduzida em seqüências de aminoácidos nos produtos polipeptídicos dos genes por máquinas macromoleculares complexas chamadas ribossomos.

- O processo de tradução é complexo, necessitando da participação de muitos RNA diferentes e moléculas protéicas.
- As moléculas de RNA transportador servem como adaptadores, mediando a interação de aminoácidos e códons no mRNA.
- O processo de tradução envolve a iniciação, o alongamento e o término de cadeias polipeptídicas e é controlado por especificações do código genético.

O CÓDIGO GENÉTICO

O código genético é um código sem superposição, com cada aminoácido mais o início e término do polipeptídeo especificados por códons do RNA compostos de três nucleotídeos.

Como ficou evidente que os genes controlam a estrutura dos polipeptídeos, a atenção concentrou-se na maneira como a seqüência dos quatro nucleotídeos diferentes no DNA pode controlar a seqüência dos 20 aminoácidos presentes nas proteínas. Com a descoberta do mRNA intermediário, surgiu a dúvida sobre como a seqüência de quatro bases presentes nas moléculas de mRNA poderiam especificar a seqüência de aminoácidos de um polipeptídeo. Qual a natureza do código genético relacionando seqüências de bases do mRNA com seqüências de aminoácidos? Obviamente, os símbolos ou letras usados no código devem ser as bases; mas o que constitui um códon, a unidade ou palavra que especifica um aminoácido ou, na verdade, um aminoacil-tRNA?

Propriedades do Código Genético: Uma Visão Geral

As principais características do código genético foram desenvolvidas durante os anos 1960. Decifrar o código foi um dos eventos mais empolgantes na história da ciência, com novas informações relatadas quase diariamente. Em meados dos anos 1960, o código genético foi amplamente esclarecido. Antes de enfocar as características específicas do código, vamos considerar suas propriedades mais importantes.

1. *O código genético é composto de trincas de nucleotídeos.* Três nucleotídeos no mRNA especificam um aminoácido no produto polipeptídico. Assim, cada códon contém três nucleotídeos.
2. *O código genético é não-superposto.* Cada nucleotídeo no mRNA pertence a apenas um códon, exceto em casos raros em que os genes se superpõem e uma seqüência de nucleotídeos é lida em duas matrizes de leitura diferentes.
3. *O código genético é sem vírgula.* Não existem vírgulas ou outras formas de pontuação dentro das regiões codificantes das moléculas de mRNA. Durante a tradução, os códons são lidos consecutivamente.

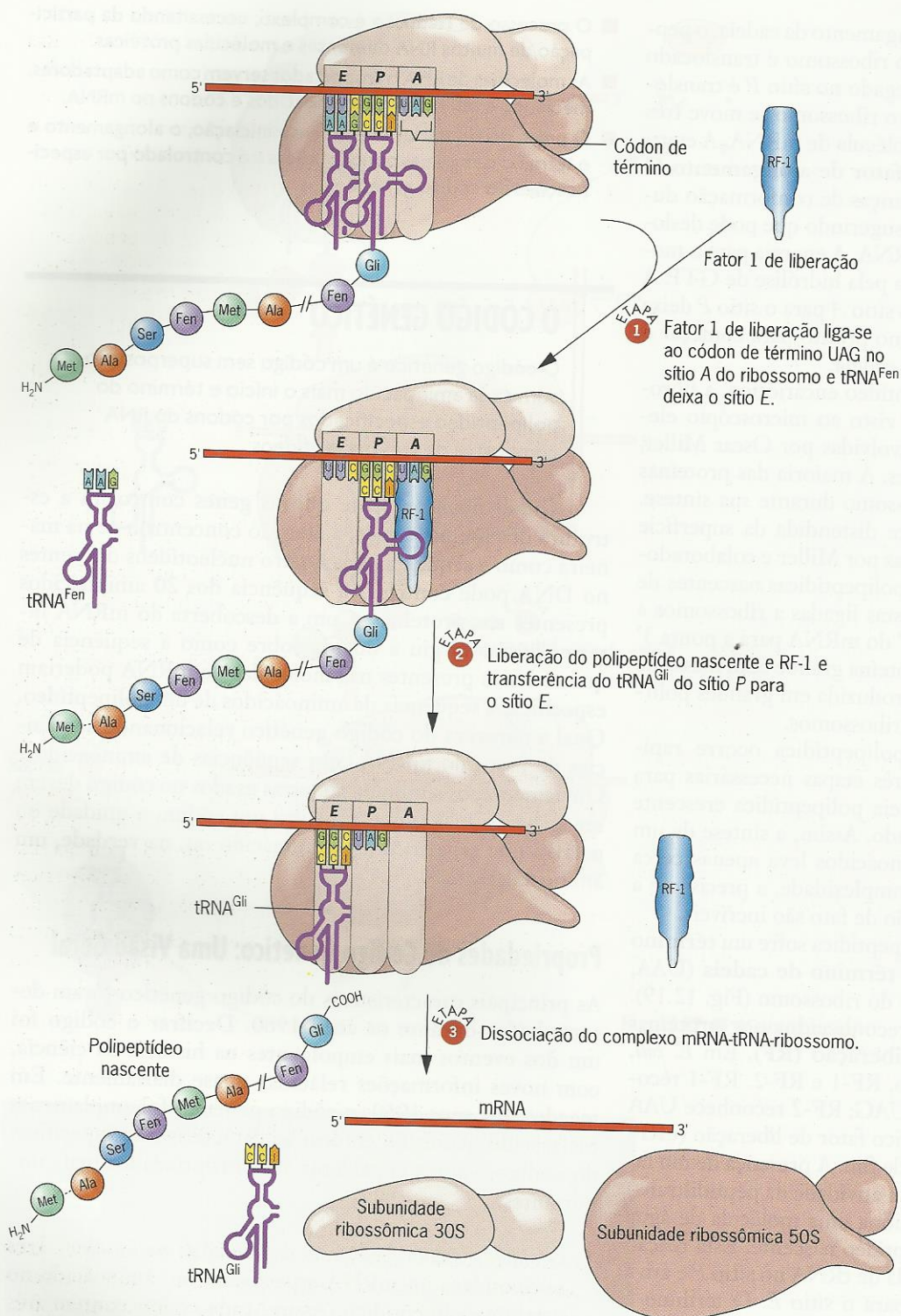


Fig. 12.19 ■ Término da cadeia polipeptídica em *E. coli*. O grupo formil de formilmetionina é removido durante a tradução.

4. O código genético é redundante. Exceto dois aminoácidos, os demais são especificados por mais de um códon.
5. O código genético é ordenado. Vários códons para um determinado aminoácido e códons para aminoácidos com propriedades químicas similares são proximamente relacionados, em geral diferindo por um único nucleotídeo.
6. O código genético contém códons de início e de fim. Códons específicos são usados para iniciar e terminar cadeias polipeptídicas.
7. O código genético é quase universal. Com pequenas exceções, os códons têm o mesmo significado em todos os organismos vivos, de vírus a humanos.

Três Nucleotídeos por Códon

Vinte aminoácidos diferentes são incorporados em polipeptídeos durante a tradução. Assim, pelo menos 20 códons diferentes devem ser formados com as quatro bases disponíveis no mRNA. Duas bases por códon resultariam em apenas 4^2 ou 16 códons possíveis, obviamente insuficientes. Três bases por códon dariam 4^3 ou 64 códons possíveis, um excesso aparente.

Em 1961, Francis Crick e colaboradores publicaram a primeira evidência forte em apoio ao **código triplo** (três nucleotídeos por códon). Crick e colaboradores fizeram uma análise genética das mutações induzidas no *locus rII* do bacteriófago T4 com a substância proflavina. A proflavina é um agente mutagênico que causa adições e deleções de um par de bases (Cap. 13). Os mutantes *rII* do fago T4 são incapazes de crescer em células de *E. coli* da linhagem K12, mas crescem como o fago tipo selvagem em células de *E. coli* linhagem B (Cap. 8). O T4 tipo selvagem cresce igualmente bem em ambas as linhagens. Crick e colaboradores isolaram revertentes induzidos por proflavina de uma mutação induzida por esta substância. Foi demonstrado que tais revertentes resultam da ocorrência de mutações adicionais em sítios vizinhos, e não de reversão da mutação original. Mutações de sítio secundário que restauram o fenótipo tipo selvagem em um organismo mutante são chamadas **mutações supressoras** porque elas cancelam, ou suprimem, os efeitos da mutação original.

Crick e colaboradores raciocinaram que se a mutação original foi uma adição ou deleção de um só par de bases, então as mutações supressoras devem ser deleções ou adições de um único par de bases, respectivamente, ocorrendo em um sítio ou sítios perto da mutação original. Se as trincas sequenciais de nucleotídeos em um mRNA especificam aminoácidos, então cada sequência de nucleotídeos pode ser reconhecida ou lida durante a tradução de três modos diferentes. Por exemplo, a sequência AAAGGGCCCTTT pode ser lida (1) AAA, GGG, CCC, TTT, (2) A, AAG, GGC, CCT, TT ou (3) AA, AGG, GCC, CTT, T. A **matriz de leitura** de um mRNA é a série de trincas de nucleotídeos que são lidos (posicionados no sítio A do ribossomo) durante a tradução. A adição ou deleção de um único par de bases irá alterar a matriz de leitura do gene e do mRNA para esta parte do gene distal à mutação. Este efeito é ilustrado na Fig. 12.20a. As mutações supressoras foram então isoladas como mutantes únicos por triagem da prole de retrocruzamentos com o tipo selvagem. Como a mutação original, foi descoberto que as mutações supressoras produzem fenótipos mutantes. Crick e colaboradores em seguida isolaram mutações supressoras induzidas por proflavina das mutações supressoras originais, e assim em diante.

Crick e colaboradores então classificaram todas as mutações isoladas em dois grupos, mais (+) e menos (-) (para adições e deleções, embora não tivessem idéia de qual grupo era cada), com base no raciocínio de que uma mutação (+) iria suprimir uma mutação (-) mas não outra mutação (+), e

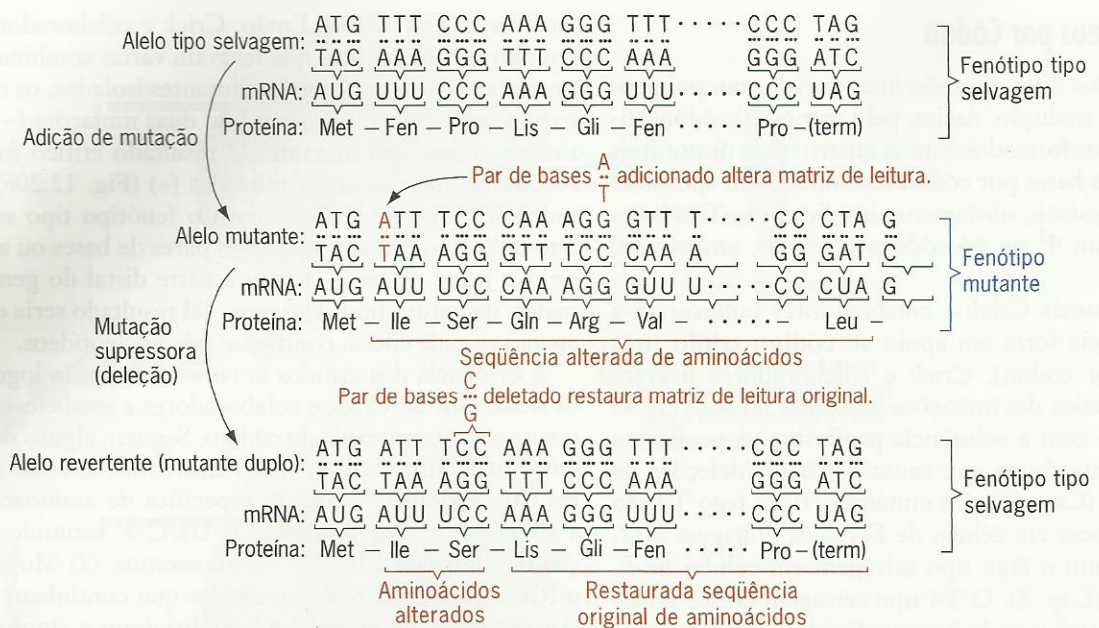
vice-versa (Fig. 12.20). Então, Crick e colaboradores construíram recombinantes que levavam várias combinações das mutações (+) e (-). Como os mutantes isolados, os recombinantes com duas mutações (+) ou duas mutações (-) sempre tinham o fenótipo mutante. O resultado crítico foi que os recombinantes com três mutações (+) (Fig. 12.20b) ou três mutações (-) em geral exibiram o fenótipo tipo selvagem. Isto indicou que a adição de três pares de bases ou a deleção de três pares de bases deixou a parte distal do gene com a matriz de leitura tipo selvagem. Tal resultado seria esperado apenas se cada códon contivesse três nucleotídeos.

A evidência dos estudos *in vitro* da tradução logo apoiou os resultados de Crick e colaboradores e estabeleceu firmemente a natureza tripla do código. Seguem alguns dos resultados mais importantes: (1) Trinucleotídeos eram suficientes para estimular a ligação específica de aminoacil-tRNA a ribossomos. Por exemplo, 5'-UUC-3' estimulou a ligação de fenilalanil-tRNA^{fen} a ribossomos. (2) Moléculas de mRNA quimicamente sintetizadas que continham sequências repetidas de dinucleotídeos dirigiram a síntese de copolímeros (grandes moléculas tipo cadeia compostas de duas subunidades diferentes) com sequências alternadas de aminoácidos. Por exemplo, quando poli(UG)_n foi usado como um mRNA artificial em um sistema de tradução *in vitro*, foi sintetizado o co-polímero repetido (cis-val)_m. (Os subscritos *n* e *m* referem-se ao número de nucleotídeos e aminoácidos nos respectivos polímeros.) (3) Em contraste, os mRNA com sequências repetidas de trinucleotídeos dirigiram a síntese de uma mistura de três homopolímeros (sendo a iniciação aleatória em tais mRNA nos sistemas *in vitro*). Por exemplo, poli(UUG)_n dirigiu a síntese de uma mistura de polileucina, policisteína e polivalina. Tais resultados são consistentes apenas com um código triplo, com suas três matrizes de leitura diferentes. Quando poli(UUG)_n é traduzida na matriz de leitura 1, UUG, UUG, é produzida uma polileucina, enquanto a tradução na matriz de leitura 2, UGU, UGU, produz policisteína, e a tradução na matriz de leitura 3, GUU, GUU, produz polivalina. Finalmente, a natureza tripla do código foi definitivamente estabelecida comparando-se as sequências de nucleotídeos de genes e mRNA com as sequências de aminoácidos de seus produtos polipeptídicos.

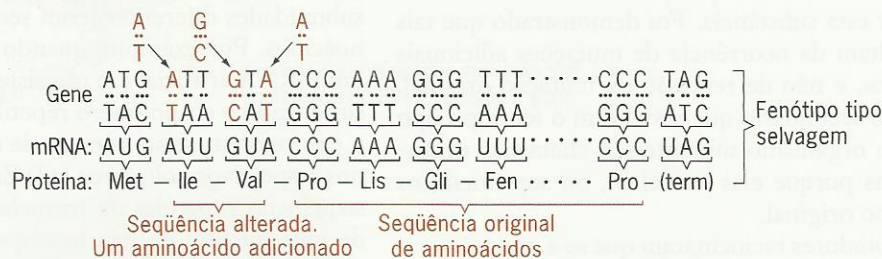
Decifrando o Código

Decifrar o código genético nos anos 1960 levou vários anos e envolveu uma intensa competição entre muitos laboratórios de pesquisa. Novas informações acumularam-se rapidamente, mas às vezes eram inconsistentes com dados anteriores. Decifrar o código foi um grande desafio.

Foram necessários cientistas para decifrar o código genético e obter respostas a várias dúvidas. (1) Que códons especificam cada um dos 20 aminoácidos? (2) Quantos dos 64 possíveis códons de trincas são utilizados? (3) Como o código é pontuado? (4) Os códons têm os mesmos significados em vírus, bactérias, plantas e animais? As respostas a estas perguntas foram obtidas primariamente dos resultados de



(a) A deleção de um único par de bases restaura a matriz de leitura alterada pela adição de um único par de bases.



(b) Recombinante contendo três adições de um par de bases tem a matriz de leitura tipo selvagem.

Fig. 12.20 ■ Evidência inicial de que o código genético é um código de trincas. Veja o texto para detalhes.

dois tipos de experimentos, ambos feitos com sistemas livres de células. O primeiro tipo de experimento envolveu a tradução de moléculas artificiais de mRNA *in vitro* e a determinação de quais dos 20 aminoácidos foram incorporados às proteínas. No segundo tipo de experimento, os ribossomos foram ativados com mini-mRNA com apenas três nucleotídeos de tamanho. Então, os pesquisadores determinaram quais aminoácil-tRNA foram estimulados a se ligar a ribossomos ativados com cada uma das mensagens de trinucleotídeos.

A década de 1960, a era do decifrar do código genético, foi uma das épocas mais empolgantes na história da biologia. Decifrar o código genético foi uma tarefa difícil e trabalhosa, e os progressos vieram em uma série de conquistas. Nós discutimos estes desenvolvimentos importantes em Um Marco na Genética: Decifrando o Código Genético. Combinando os resultados dos experimentos de tradução *in vitro* feitos com mRNA sintéticos e análises de ligação de trinucleotídeos, Marshall Nirenberg, Severo Ochoa, H. Ghorband Khorana, Philip Leder e seus colaboradores esclareceram o significado de todos os 64 códons de trincas (Quadro 12.1). Nirenberg e Khorana compartilharam o Prêmio Nobel de

1968 de Fisiologia ou Medicina por seu trabalho no código com Robert Holley, que determinou a sequência completa de nucleotídeos do tRNA de alanina de levedura. Ochoa já tinha recebido o Prêmio Nobel de 1959 por sua descoberta da RNA-polimerase.

Códons de Início e Término

O código genético também fornece uma pontuação da informação genética em nível de tradução. Tanto em procariontes quanto em eucariontes, o códon AUG é usado para iniciar as cadeias polipeptídicas (Quadro 12.1). Em casos raros, GUG é usado como um códon de início. Em ambos os casos, o códon de iniciação é reconhecido por um tRNA iniciador, o tRNA^{Met} em procariontes e tRNA_i^{Met} em eucariontes. Em procariontes, um códon AUG deve seguir uma sequência apropriada de nucleotídeos, a sequência Shine-Delgarno, no segmento não traduzido a 5' da molécula de mRNA para servir como códon iniciador da tradução. Em eucariontes, o códon deve ser o primeiro AUG encontrado pelo ribossomo

QUADRO 12.1

O Código Genético^a

		Segunda letra				
		U	C	A	G	
Primeira (5') letra	U	UUU } Fen UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tir UAC } UAA Ocre (finalizador) UAG Âmbar (finalizador)	UGU } Cis UGC } UGA Opala (finalizador) UGG Trp	U C A G
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }	U C A G
	A	AUU } AUC } Ile AUA } AUG Met (iniciador)	ACU } ACC } Tre ACA } ACG }	AAU } Asn AAC } AAA } Lis AAG }	AGU } Ser AGC } AGA } Arg AGG }	U C A G
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gli GGA } GGG }	U C A G

^aCada sequência de trinca de nucleotídeos ou códon refere-se à sequência de nucleotídeos no mRNA (não no DNA) que especifica a incorporação do aminoácido indicado ou polipeptídeo de término de cadeia.

■ = Códon de início de cadeia polipeptídica
■ = Códon de término de cadeia polipeptídica

à medida que ele percorre a partir da ponta 5' da molécula de mRNA. Em posições internas, AUG é reconhecido por tRNA^{Met}, e GUG é reconhecido por um tRNA de valina.

Três códons, UAG, UAA e UGA, especificam término de cadeia polipeptídica (Quadro 12.1). Devido a um evento histórico puramente frívolo, os códons UAG, UAA e UGA às vezes são chamados finalizadores *âmbar*, *ocre* e *opala*, respectivamente (veja Enfoque Técnico: O Códon "Âmbar"). Estes códons são reconhecidos por fatores de liberação protéicos, e não por tRNA. Os procariontes contêm dois fatores de liberação, RF-1 e RF-2. RF-1 termina polipeptídeos em resposta a códons UAA e UAG, enquanto RF-2 causa o término nos códons UAA e UGA. Eucariontes contêm um único fator de liberação que reconhece todos os três códons de término.

Um Código Redundante e Ordenado

Todos os aminoácidos exceto metionina e triptofano são especificados por mais de um códon (Quadro 12.1). Dos três

aminoácidos leucina, serina e arginina, cada um é especificado por seis códons diferentes. Isoleucina tem três códons. Os outros aminoácidos têm, cada um, dois ou quatro códons. A ocorrência de mais de um códon por aminoácido é chamada de **redundância**. A redundância no código genético não é aleatória, mas sim altamente ordenada. Na maioria dos casos, os vários códons que especificam um determinado aminoácido diferem por apenas uma base, a terceira base, ou 3', do códon. A redundância é primariamente de dois tipos. (1) Redundância parcial ocorre quando a terceira base pode ser uma de duas pirimidinas (U ou C) ou, alternativamente, uma de duas purinas (A ou G). Com a redundância parcial, a mudança da terceira base de uma purina para uma pirimidina, ou vice-versa, mudará o aminoácido especificado pelo códon. (2) No caso de redundância completa, qualquer uma das quatro bases pode estar presente na terceira posição no códon, e o códon ainda especificará o mesmo aminoácido. Por exemplo, valina é codificada por GUU, GUC, GUA e GUG (Quadro 12.1).



ENFOQUE TÉCNICO: O Códon “Âmbar”

Como discutido no texto, mutações que produzem códons UAG *in-frame**, que especificam término da cadeia polipeptídica, dentro das regiões codificantes dos mRNA são chamadas de mutações *âmbar*. Elas geram produtos gênicos truncados – polipeptídeos sem seus terminos carboxila normais. Os mutantes *âmbar*, bactérias ou bacteriófagos que têm mutações *âmbar*, demonstraram-se extremamente valiosos na análise da estrutura e função gênica. Mas por que são chamados mutantes *âmbar*? Todos têm a cor *âmbar*?

A maioria dos termos científicos são descritivos do componente, do processo ou do evento que representam. Entretanto, às vezes um termo frívolo é introduzido e torna-se estabelecido na ciência. O uso do termo *âmbar* para se referir ao códon de término UAG é um exemplo.

Em 1960, Harris Bernstein era um estudante que trabalhava com o fungo *N. crassa* no laboratório de Herschel Mitchell no California Institute of Technology em Pasadena. Em uma tarde, após trabalhar o dia inteiro no laboratório, Harris andava no corredor para o “laboratório de fagos”, como era chamado o laboratório de pesquisa de Delbrück, pensando em convidar seus amigos Charley Steinberg e Dick Epstein para irem ao cinema. Sem que Harris soubesse, Charley e Dick tinham acabado de começar um experimento grande e trabalhoso. Eles estavam tentando isolar mutantes do bacteriófago T4 que pudessem

crescer na linhagem C de *E. coli* mas não crescer na linhagem B de *E. coli*.

A triagem de tais mutantes envolve transferências repetidas do fago de uma placa para duas outras placas contendo linhagens hospedeiras diferentes. Cada fago mutante potencial deve ser individualmente transferido. Na época, as transferências eram feitas com agulhas inoculantes flambadas entre as transferências, um processo extremamente trabalhoso. Assim, Steinberg e Epstein disseram a Bernstein para esquecer o filme e ficar para ajudá-los em seu experimento. Harris concordou em ficar e ajudar na caça ao mutante, mas só se Steinberg e Epstein concordassem em dar ao mutante o nome de sua mãe, Hannah Bernstein, se o experimento tivesse sucesso.

Quando Bernstein chegou ao laboratório na manhã seguinte, os mutantes desejados haviam sido identificados. Além disso, eles já haviam sido denominados mutantes *ambar*, que é a tradução para o inglês da palavra alemã *Bernstein*. Entretanto, há uma discordância em um ponto. Dick Epstein, que criou o nome mutações *âmbar*, disse que o nome foi dado em homenagem a Harris Bernstein, e não à mãe de Harris.

Quando se demonstrou que tais mutações resultam em códons UAG de término de cadeia dentro das regiões codificantes dos mRNA, o termo *âmbar* foi extrapolado para o códon de término também. Posteriormente, os códons de término UAA e UGA foram denominados *ocre* e *opala* de modo a existirem termos análogos para todos os três códons de término. Assim, embora o termo *âmbar* tivesse sido introduzido de maneira frívola, hoje ele é um componente estabelecido do vocabulário dos geneticistas.

*N.T.: *In-frame* é a manutenção de uma matriz de leitura de uma região codificante após uma mutação, ou ligação de duas moléculas de DNA.

Cientistas especularam que a **ordem** no código genético tinha evoluído como um modo de minimizar a letalidade mutacional. Muitas substituições de bases na terceira posição dos códons não trocam o aminoácido especificado pelo códon. Além disso, aminoácidos com propriedades químicas similares (tais como leucina, isoleucina e valina) têm códons que diferem uns dos outros por apenas uma base. Assim, muitas substituições de um só par de bases resultarão na substituição de um aminoácido por outro com propriedades químicas muito similares (por exemplo, valina por isoleucina). Na maioria dos casos, substituições conservadoras deste tipo geram produtos gênicos ativos, que minimizam os efeitos das mutações.

Um Código Quase Universal

Grandes quantidades de informações hoje estão disponíveis a partir de estudos *in vitro*, de substituições de aminoácidos devidas a mutações e de ácidos nucleicos correlatos ou sequenciamento de polipeptídeos, que permitem uma compa-

ração do significado dos 64 códons em espécies diferentes. Todos estes dados indicam que o código genético é quase **universal**, isto é, os códons têm o mesmo significado, com pequenas exceções, em todas as espécies.

As exceções mais importantes à universalidade do códon ocorrem em mitocôndrias de mamíferos, leveduras e várias outras espécies. As mitocôndrias têm seus próprios cromossomos e maquinaria de síntese de proteínas (Cap. 19). Embora os sistemas mitocondrial e citoplasmático sejam similares, existem algumas diferenças. Nas mitocôndrias de humanos e outros mamíferos, (1) UGA especifica triptofano em vez de término de cadeia, (2) AUA é um códon de metionina, e não um códon de isoleucina, e (3) AGA e AGG são códons de término de cadeia, e não códons de arginina. Os outros 60 códons têm o mesmo significado em mitocôndrias de mamíferos que nos mRNA nucleares (Quadro 12.1). Existem também diferenças raras no significado de códons nas mitocôndrias de outras espécies e em transcritos nucleares de alguns protozoários. Entretanto, como tais exceções são raras, o código genético deve ser considerado como quase universal.

PONTOS IMPORTANTES

- Cada um dos 20 aminoácidos em proteínas é especificado por uma ou mais trincas de nucleotídeos no mRNA.
- Das 64 trincas possíveis, considerando-se as quatro bases no mRNA, 61 especificam aminoácidos e 3 indicam término de cadeia.
- O código é sem superposição, com cada nucleotídeo sendo parte de um único códon, é redundante, com a maioria dos aminoácidos sendo especificados por dois ou quatro códons, e ordenado, com aminoácidos similares especificados por códons correlatos.
- O código genético é quase universal; com pequenas exceções, as 64 trincas têm o mesmo significado em todos os organismos.

INTERAÇÕES CÓDON-tRNA

Os códons nas moléculas de mRNA são reconhecidos por aminoacil-tRNA durante a tradução.

A tradução de uma sequência de nucleotídeos no mRNA na sequência correta de aminoácidos no produto polipeptídico requer o reconhecimento preciso de códons pelos aminoacil-tRNA. Devido à redundância do código genético, ou vários tRNA diferentes devem reconhecer códons diferentes que especificam um determinado aminoácido ou o anticódon de um determinado tRNA deve ser capaz de fazer pares de bases com vários códons diferentes. De fato, ocorrem ambos os fenômenos. Existem vários tRNA para alguns aminoácidos, e alguns tRNA reconhecem mais de um códon.

Reconhecimento dos Códon por tRNA: A Hipótese da Oscilação

As pontes de hidrogênio entre as bases nos anticódons dos tRNA e os códons dos mRNA seguem estritamente as regras de pareamento de bases apenas para as primeiras duas bases do códon. O pareamento de bases que envolve a terceira base do códon é menos restritivo, permitindo o que Crick chamou **oscilação** (*wobble*) neste ponto. Com base em distâncias moleculares e considerações estéricas (estrutura tridimensional), Crick propôs que a oscilação permitiria vários tipos de pareamento de bases, mas não todos os tipos, na terceira base do códon – interação códon-anticódon. Sua proposta foi amplamente apoiada por dados experimentais. O Quadro 12.2 mostra o pareamento de bases previsto pela hipótese da oscilação de Crick.

A hipótese da oscilação prevê a existência de pelo menos dois tRNA para cada aminoácido com códons que apresentem redundância completa, e isto foi confirmado. A hipótese da oscilação também prevê a ocorrência de três tRNA para os seis códons de serina. Três tRNA de serina foram carac-

QUADRO 12.2

Pareamento de Bases Entre a Base 5' dos Anticódons dos tRNA e a Base 3' dos Códon dos mRNA de acordo com a Hipótese da Oscilação

Base no Anticódon	Base no Códon
G	U ou C
C	G
A	U
U	A ou G
I	A, U ou C

terizados: (1) tRNA^{Ser1} (anticódon AGG) liga-se a códons UCU e UCC, (2) tRNA^{Ser2} (anticódon AGU) liga-se a códons UCA e UCG e (3) tRNA^{Ser3} (anticódon UCG) liga-se a códons AGU e AGC. Estas especificidades foram verificadas pela ligação estimulada por trinucleotídeos de aminoacil-tRNA purificados a ribossomos *in vitro*.

Finalmente, vários tRNA contêm a base inosina, que é feita a partir da purina hipoxantina. Inosina é produzida por modificação pós-transcricional de adenosina. A hipótese de Crick da oscilação prevê que, quando a inosina está presente na ponta 5' de um anticódon (a posição de oscilação), ela iria parear com uracil, citosina ou adenina no códon. De fato, alanil-tRNA purificado contendo inosina (I) na posição 5' do anticódon (veja Fig. 12.11) liga-se a ribossomos ativados com trinucleotídeos GCU, GCC ou GCA (Fig. 12.21). O mesmo resultado foi obtido com outros tRNA purificados com inosina na posição 5' do anticódon. Assim, a hipótese de Crick da oscilação explica perfeitamente as relações entre tRNA e códons considerando-se a redundância, mas ordenação, do código genético.

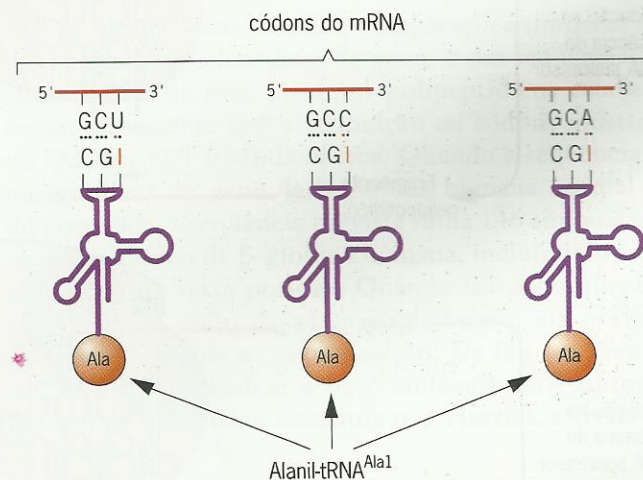


Fig. 12.21 ■ Pareamento de bases entre o anticódon de alanil-tRNA^{Ala1} e os códons GCU, GCC e GCA do mRNA de acordo com a hipótese de Crick da oscilação. Análises de ligação de ribossomos ativados por trinucleotídeos mostraram que o alanil-tRNA^{Ala1} de fato faz pareamento de bases com todos os três códons.

Mutações Supressoras que Produzem tRNA com Reconhecimento Alterado de Códon

Mesmo excluindo as mitocôndrias, o código genético não é absolutamente universal. Pequenas variações no reconhecimento e na tradução de códon estão bem documentadas. Em *E. coli* e leveduras, por exemplo, algumas mutações nos genes de tRNA alteram os anticódons e portanto os códon reconhecidos pelos tRNA mutantes. Tais mutações foram inicialmente detectadas como *mutações supressoras*, substituições de nucleotídeos que suprimem os efeitos de outras mutações. Foi subsequentemente mostrado que as mutações supressoras ocorrem em genes de tRNA. Mui-

tas dessas mutações supressoras mudaram os anticódons de tRNA alterados.

Os exemplos mais conhecidos de mutações supressoras que alteram a especificidade do tRNA são as que suprimem o finalizador de cadeia UAG dentro das seqüências codificantes de genes. Tais mutações, chamadas *âmbar* (veja Enfoque Técnico: O Códon “Âmbar”), resultam na síntese de polipeptídeos truncados. Mutações que produzem trinças de término de cadeia dentro dos genes ficaram conhecidas como **mutações sem sentido**, em contraste com as **mutações de sentido trocado**, que mudam uma trinca de modo que ela especifique um aminoácido diferente. Um gene que contém uma mutação de sentido trocado codifica um polipeptídeo completo, mas com

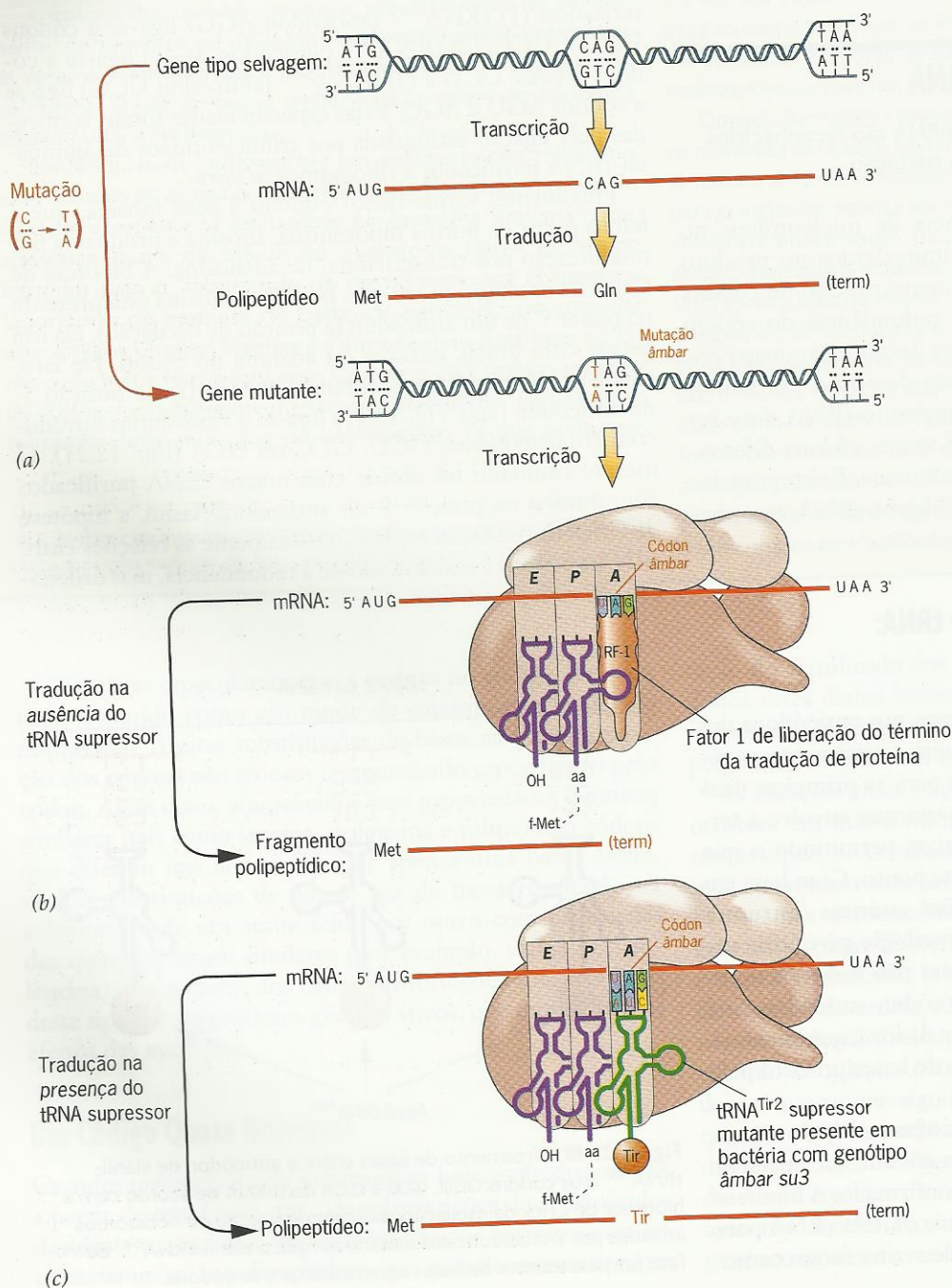


Fig. 12.22 ■ (a) A formação de uma mutação *âmbar* (UAG) de término de cadeia. (b) Seu efeito no produto gênico polipeptídico na ausência de um tRNA supressor e (c) na presença de um tRNA supressor. A mutação *âmbar* mostra mudanças de um códon de glutamina CAG (Gln) para um códon UAG de término de cadeia. O polipeptídeo contendo a tirosina inserida pelo tRNA supressor pode ou não ser funcional. Entretanto, supressão do fenótipo mutante irá ocorrer apenas quando o polipeptídeo é funcional.

uma substituição de aminoácidos no produto gênico polipeptídico. Uma mutação sem sentido resulta em um polipeptídeo truncado, com o tamanho da cadeia dependendo da posição da mutação dentro do gene. Mutações sem sentido frequentemente resultam de substituições de um único par de bases, como ilustrado na Fig. 12.22a. Os fragmentos polipeptídicos produzidos de genes contendo mutações sem sentido (Fig. 12.22b) em geral são totalmente não-funcionais.

Foi demonstrado que supressão de mutações sem sentido resulta de mutações em genes de tRNA que fazem com que tRNA mutantes reconheçam os códons de término (UAG, UAA ou UGA), embora com eficiências variáveis. Esses tRNA mutantes são chamados **tRNA supressores**. Quando o tRNA supressor *âmbar* (UAG) produzido pela mutação *âmbar* *su3* em *E. coli* foi sequenciado, viu-se que ele tem um anticódon alterado. Essa mutação particular no supressor *âmbar* ocorre no gene de tRNA^{Tir2} (um dos dois genes de tRNA de tirosina em *E. coli*). Foi demonstrado que o anticódon do tipo selvagem tRNA^{Tir2} (não supressor) é 5'-G'UA-3' (em que G' é um derivado de guanina). O anticódon do tRNA^{Tir2} mutante (supressor) é 5'-CUA-3'. Devido à substituição de uma única base, o anticódon do supressor tRNA^{Tir2} faz pares de bases com o códon *âmbar* 5'-UAG-3' (lembre que o pareamento de bases sempre envolve filamentos de polaridade inversa); isto é,

mRNA: 5'-UAG-3' (códon)
tRNA: 3'-AUC-5' (anticódon)

Portanto, os tRNA supressores permitem que sejam produzidos polipeptídeos completos a partir de mRNA contendo códons de término dentro dos genes (Fig. 12.22c). Tais polipeptídeos serão funcionais se o aminoácido inserido pelo tRNA supressor não alterar significativamente as propriedades químicas da proteína.

PONTOS IMPORTANTES

- A hipótese da oscilação explica como um único tRNA pode responder a dois ou mais códons.
- Algumas mutações supressoras alteram os anticódons de tRNA de modo que tRNA mutantes reconhecem os códons de término de cadeia e inserem aminoácidos em resposta à sua presença nas moléculas de mRNA.

CONFIRMAÇÃO *IN VIVO* DA NATUREZA DO CÓDIGO GENÉTICO

Comparações entre as seqüências de nucleotídeos de genes e as seqüências de aminoácidos de seus produtos polipeptídicos mostraram que as atribuições de códons deduzidas de estudos *in vitro* são usadas durante a síntese de proteínas em células vivas.

As atribuições dos códons mostradas no Quadro 12.1 foram inicialmente baseadas nos resultados de tra-

dução *in vitro* e estudos de ligação de aminoacil-tRNA estimulada por trinucleotídeos. A fonte dessas atribuições levantou uma questão óbvia. As atribuições baseadas em experimentos *in vitro* são válidas *in vivo*? Os projetos recentes de sequenciamento do genoma demonstraram que tais atribuições de códons são corretas para a síntese de proteínas *in vivo* para muitas espécies ou mesmo para todas. Quando as substituições de aminoácidos que resultam de mutações induzidas por mutágenos químicos com efeitos específicos (Cap. 13) foram determinadas por sequenciamento de aminoácidos, as substituições eram quase sempre consistentes com as atribuições de códons fornecidas no Quadro 12.1 e o efeito conhecido do mutágeno.

Mais convincentemente, quando as seqüências de nucleotídeos dos genes ou mRNA foram determinadas e comparadas com as seqüências de aminoácidos de polipeptídeos codificados por estes genes ou mRNA, as correlações observadas foram as previstas pelas atribuições de códons mostradas no Quadro 12.1. Esta correlação foi primeiro demonstrada para o bacteriófago MS2 por Walter Fiers e colegas em 1972. Fiers e colaboradores determinaram a seqüência de aminoácidos da capa de proteínas do fago MS2 e a seqüência de nucleotídeos do gene MS2 que codifica a capa de proteínas (veja Fig. 14.9). O fago MS2 estoca sua informação genética em RNA unifilar; seu cromossomo é equivalente a uma molécula de mRNA em organismos com genomas de DNA. Fiers e colaboradores então compararam a seqüência de nucleotídeos do gene para a capa de proteína com a seqüência de aminoácidos da capa de polipeptídeo. A seqüência de aminoácidos da capa de proteína era exatamente a seqüência prevista pela seqüência de nucleotídeos do gene para a capa de proteína e as atribuições de códons mostradas no Quadro 12.1.

Nas últimas duas décadas, comparações similares das seqüências de nucleotídeos de genes e das seqüências de aminoácidos de seus produtos polipeptídicos estabeleceram claramente que a atribuição de códons mostrada no Quadro 12.1 é válida *in vivo*. Quando a seqüência de nucleotídeos do gene da β -globina humana normal foi determinada, a seqüência prevista tinha 146 aminoácidos no polipeptídeo de β -globina humana, incluindo o ácido glutâmico na sexta posição. Quando foi determinada a seqüência do alelo da anemia falciforme, ela previa a presença de valina na sexta posição. De fato, os cientistas hoje compreendem a base molecular dos sintomas da anemia falciforme descritos por Herrick e Irons em 1910.

PONTOS IMPORTANTES

- As comparações das seqüências de nucleotídeos dos genes com as seqüências de aminoácidos de seus produtos polipeptídicos verificaram as atribuições de códons deduzidas de estudos *in vitro*.



UM MARCO NA GENÉTICA: Decifrando o Código Genético

Quando os cientistas começaram a decifrar o código genético em 1960, eles sabiam que as quatro letras do alfabeto genético, os quatro nucleotídeos no RNA, de algum modo tinham que especificar os 20 aminoácidos nos produtos polipeptídicos dos genes mais a iniciação e o término das cadeias polipeptídicas. Um código duplo, dois nucleotídeos por aminoácido, não daria informações suficientes ($4^2 = 16$ códons). Entretanto, um código de trincas parecia fornecer informações demais ($4^3 = 64$ códons). Então, em 1961, Francis Crick e colaboradores deram fortes evidências de um código de trincas (veja Fig. 12.20). Mas que códons de trincas especificavam quais aminoácidos?

A primeira conquista surgiu em 1961 quando Marshall Nirenberg (ganhador do Prêmio Nobel de 1968) e J. Heinrich Matthaei demonstraram que moléculas sintéticas de RNA poderiam ser usadas como mRNA artificiais para a síntese direta de proteínas *in vitro*.¹ Quando ribossomos, aminoacil-tRNA e fatores solúveis necessários para a tradução são purificados livres de mRNA naturais, estes componentes podem ser combinados *in vitro* e estimulados a sintetizar polipeptídeos pela adição de moléculas de RNA sintetizadas quimicamente. Se tais moléculas sintéticas de mRNA são de conteúdo nucleotídico conhecido, a composição de aminoácidos dos polipeptídeos resultantes pode ser usada para deduzir quais códons especificam quais aminoácidos.

A primeira atribuição de códons (UUU para fenilalanina) foi feita quando Nirenberg e Matthaei demonstraram que o ácido poliurídico [poli(U) = (U)_n] dirige a síntese de polifenilalanina [(fenilalanina)_n]. Eles usaram fenilalanina marcada radioativamente (¹⁴C-fenilalanina) como um substrato em um sistema de tradução *in vitro* com poli(U), poli(C) e poli(A) como RNA mensageiros artificiais. A ¹⁴C-fenilalanina foi incorporada em polifenilalanina apenas quando poli(U) foi usado como o mRNA (Quadro 1). Quando cada um dos 17 outros aminoácidos marcados foi usado no sistema de tradução estimulado por poli(U), nenhuma radioatividade foi incorporada a polipeptídeos. Devido à natureza tripla do código genético, Nirenberg e Matthaei concluíram que UUU deve ser um códon para fenilalanina. Logo depois, foi demonstrado que poli(A) e poli(C) codificam polilisina e poliprolina, respectivamente, permitindo que Nirenberg e colaboradores atribuísem o códon AAA a lisina e CCC a prolina. Poli(G) não funciona como um mRNA nos sistemas de tradução *in vitro* porque o pareamento de pares de bases entre guaninas resulta em estruturas trifilamentares complexas.

Pesquisadores nos laboratórios de Nirenberg e Severo Ochoa (ganhador do Prêmio Nobel de 1959) ampliaram o uso de mRNA artificiais para co-polímeros sintéticos com seqüências aleatórias. Por exemplo, Nirenberg e colaboradores sintetizaram um co-po-

QUADRO 1

Incorporação de ¹⁴C-Fenilalanina a Polifenilalanina em um Sistema de Tradução *In Vitro* Ativado com Homopolímeros de RNA Sintético^a

mRNA Sintético	Radioatividade Incorporada (contagens por minuto)
Nenhum	44
Poli(U)	39.800
Poli(A)	50
Poli(C)	38
Poli(I) ^b	57

^aDados de Nirenberg e Matthaei, 1961. *Proc. Natl. Acad. Sci. USA* 47: 1588-1602.

^bPoli (I) é ácido poliinosínico, que contém a purina hipoxantina. A hipoxantina é como a guanina, pois faz par de base com citosina.

límero aleatório contendo quantidades aproximadamente iguais de adenina e citosina e o usaram como um mRNA artificial em seu sistema de tradução *in vitro*.² Um co-polímero aleatório AC com quantidades iguais de A e C irá conter 12,5% ($\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$) de cada um dos oito possíveis códons: AAA, AAC, ACA, CAA, CCA, CAC, ACC e CCC. Nirenberg e colaboradores observaram que poli(AC) dirigia a incorporação de seis aminoácidos – asparagina, glutamina, histidina, lisina, prolina e treonina – em polipeptídeos. Como eles já sabiam que AAA e CCC eram códons de lisina e prolina, seus resultados indicaram que os códons compostos de dois A mais um C e dois C mais um A especificam asparagina, glutamina, histidina e treonina. Quantidades aproximadamente iguais de asparagina, glutamina, histidina e lisina eram incorporadas, mas treonina e prolina eram incorporadas em cerca do dobro da quantidade de cada um dos outros quatro aminoácidos. Este resultado indicou que co-polímeros aleatórios AC podem conter o dobro de códons de treonina e prolina que os códons de asparagina. De fato, existem dois códons de treonina (ACC e ACA) e dois códons de prolina (CCC e CCA) em co-polímeros aleatórios AC, mas apenas um códon para cada um dos outros quatro aminoácidos – por exemplo, asparagina (AAC).

Variando a composição de nucleotídeos de co-polímeros aleatórios, Ochoa e colaboradores alteraram as frequências relativas dos oito códons e procuraram correlações com as frequências de aminoácidos nos polipeptídeos sintetizados em resposta aos co-polímeros.³ Por exemplo, eles sintetizaram co-polímeros aleatórios contendo A e C nas proporções 5:1 e 1:5. Quando esses co-polímeros foram usados como mRNA *in vitro*, os mesmos seis

¹Nirenberg, M. W., e J. H. Matthaei. 1961. The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* 47: 1588-1602.

²Nirenberg, M. W., O. W. Jones, P. Leder, B. F. C. Clark, W. S. Sly e S. Pestka. 1963. On the coding of genetic information. *Cold Spring Harbor Symp. Quant. Biol.* 28: 549-557.

³Speyer, J. F., P. Lengyel, C. Basilio e S. Ochoa. 1962. Synthetic polynucleotides and the amino acid code, IV. *Proc. Natl. Acad. Sci. U.S.A.* 48: 441-448.



UM MARCO NA GENÉTICA (continuação)

Trinucleotídeo de mini-mRNA



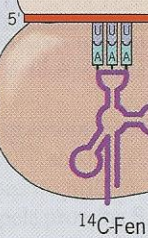
¹⁴C-Fen

Fenilalanil-tRNA^{Fen}
marcado com ¹⁴C

Subunidade
ribossômica 30S

Subunidade ribossômica 50S

O trinucleotídeo ativa o
ribossomo e estimula a
ligação do aminoacil-tRNA.



¹⁴C-Fen

Complexo trinucleotídeo/fenilalanil-tRNA^{Fen}/ribossomo

Fig. 1 ■ Estimulação da ligação de aminoacil-tRNA a ribossomos por trinucleotídeos sintéticos mini-mRNA. Os resultados destas análises de ligação de ribossomos ativados por trinucleotídeos ajudaram os cientistas a decifrar o código genético.

aminoácidos foram incorporados como nos experimentos de Nirenberg com 1A:1C de co-polímeros aleatórios, mas suas frequências relativas nos produtos polipeptídicos foram muito diferentes. Com um co-polímero 1A:5C, prolina e lisina representam 60,8% e 1,2%, respectivamente, dos aminoácidos incorporados, em concordância com a demonstração anterior de que CCC e AAA eram códons de prolina e lisina. Os valores foram revertidos quando foi usado o co-polímero 5A:1C; 51% e 3,8% dos aminoácidos incorporados eram lisina e prolina, respectivamente. Os resultados de Ochoa indicaram que os códons de glutamina e asparagina continham 2A e 1C, enquanto o códon de histidina continha 1A e 2C, e assim em diante.

Uma grande conquista ocorreu quando H. Gobind Khorana (ganador do Prêmio Nobel de 1968) e colaboradores desenvolveram um procedimento pelo qual co-polímeros com seqüências conhecidas repetidas de di-, tri- e tetranucleotídeos podiam ser

sintetizados.⁴ Como as seqüências de códons nesses co-polímeros repetidos eram fixas, seu uso como mRNA sintéticos deu resultados mais facilmente interpretados que os obtidos com co-polímeros aleatórios. Por exemplo, uma molécula de RNA com uma seqüência de dinucleotídeo UG repetida dirigia a síntese de polipeptídeos contendo cisteínas e valinas alternadas, como mostrado a seguir.

5'-UGU GUG UGU GUG UGU GUG-3'

H₂N—Cis—Val—Cis—Val—Cis—Val—COOH

Como as trincas UGU e GUG se alternam no poli(UG)_n, estes dois códons devem especificar cisteína e valina, mas o resultado não nos diz qual códon especifica qual aminoácido. Em contraste, um polímero de trinucleotídeos repetidos UUG dirige a síntese de uma mistura de polileucina, policisteína e polivalina.

mRNA sintético

5'-UUG UUG UUG UUG UUG UUG UUG

Matriz de
leitura 1

Matriz de
leitura 2

Matriz de
leitura 3

Produtos polipeptídicos

Polileucina
Policisteína
Polivalina

Leu—Leu—Leu—Leu—Leu—Leu—Leu
Cis—Cis—Cis—Cis—Cis—Cis—Cis
Val—Val—Val—Val—Val—Val—Val

Como o início da tradução ocorre aleatoriamente nestes sistemas aleatórios *in vitro*, alguns ribossomos irão traduzir tais polímeros como UUG, UUG, UUG, e assim em diante, enquanto outros irão traduzi-los como UGU, UGU, UGU,..., e ainda outros como GUU, GUU... Assim, estes três códons devem especificar leucina, cisteína e valina. Como o produto poli(UG)_n era um co-polímero cisteína-valina, um dos códons adicionais em poli(UUG)_n, seja UUG ou GUU, deve ser um códon de leucina. Analisando os aminoácidos incorporados em resposta a diferentes co-polímeros repetidos de di-, tri- e tetranucleotídeos, Khorana e colaboradores foram capazes de atribuir muitos destes códons a aminoácidos específicos. Por exemplo, UGU especifica cisteína, UUG leucina e GUU valina.

Informações adicionais sobre a natureza do código genético foram obtidas analisando-se a ligação de aminoacil-tRNA a ribossomos ativados com pequenos oligômeros de RNA. Quando um sistema de tradução *in vitro* é ativado com poli(U), apenas um aminoacil-tRNA, fenilalanil-tRNA^{Fen}, liga-se ao ribossomo. Esta

⁴Khorana, H. G., H. Büchi, H. Ghosh, N. Gupta, T. M. Jacob, H. Kössel, R. Morgan, S. A. Narang, E. Ohtsuka e R. D. Wells. 1966. Polynucleotide synthesis and the genetic code. *Cold Spring Harbor Symp. Quant. Biol.* 31: 39-49.



UM MARCO NA GENÉTICA (continuação)

especificidade existe porque o códon do mRNA é parte do sítio de ligação, e a ligação envolve o pareamento de bases entre o códon e o anticódon do tRNA. Em 1964, Nirenberg e Philip Leder desenvolveram uma análise para a ligação de aminoacil-tRNA a ribossomos ativados com *trinucleotídeos*, mini-mRNA com apenas três nucleotídeos de tamanho.⁵ Nirenberg e Leder sintetizaram trinucleotídeos de sequência conhecida e testaram sua capacidade de estimular a ligação de aminoacil-tRNA específicos a ribossomos. Eles analisaram a capacidade de cada trinucleotídeo de servir como um mini-mRNA usando aminoácidos marcados para detectar a formação de complexos trinucleotídeo-aminoacil-tRNA-ribossomo (Fig. 1). Por exemplo, os trinucleotídeos 5'-UUU-3' e 5'-UUC-3' estimulavam a ligação de fenilalanil-tRNA^{Fen} a ribossomos, indicando que UUU e UUC são ambos códons de fenilalanina.

Combinando os resultados de análises de ligação de trinucleotídeos e experimentos de tradução *in vitro* feitos com mRNA

⁵Nirenberg, M., e P. Leder. 1964. RNA codewords and protein synthesis: The effect of trinucleotides upon the binding of sRNA to ribosomes. *Science* 145: 1399-1407.

sintéticos, Nirenberg, Ochoa, Khorana, Leder e outros foram capazes de decifrar o significado dos 64 códons (veja Quadro 12.1). Estas atribuições de códons hoje estão firmemente estabelecidas, apoiadas por dados definitivos de estudos tanto *in vitro* quanto *in vivo*.

QUESTÕES PARA DISCUSSÕES

1. O código genético é amplamente o mesmo, isto é, cada códon triplo tem o mesmo significado em todos os organismos – vírus, bactérias, fungos, plantas e animais. Esta quase universalidade do código tem alguma implicação na evolução e na origem das espécies? Caso sim, qual? Caso não, esta quase universalidade do código pode ser explicada?
2. Se existirem organismos vivos que estocam informações genéticas em ácidos nucleicos e usam proteínas para catálise em outros planetas, você acha que usariam um código genético idêntico ao nosso? Por que ou por que não?

Exercícios Básicos Ilustram a análise genética básica.

1. O polipeptídeo de β -globina humana tem 146 aminoácidos de tamanho. Qual o tamanho da parte codificante do mRNA de β -globina humana?

Resposta: Cada aminoácido é especificado por um códon contendo três nucleotídeos. Portanto, os 146 aminoácidos na β -globina serão especificados por 438 (146×3) nucleotídeos. Entretanto, um códon de fim deve estar presente no final da sequência codificante, levando o tamanho para $438 + 3 = 441$ nucleotídeos. No caso de β -globina e muitas outras proteínas, a metionina amino-terminal (especificada pelo códon de iniciação AUG) é removida da β -globina durante a síntese. A adição do códon de iniciação aumenta a sequência codificante do mRNA de β -globina para 444 nucleotídeos ($441 + 3$).

2. Se o segmento codificante de um mRNA com a sequência 5'-AUGUUUCCCAAAGGG-3' é traduzida, que sequência de aminoácidos será produzida?

Resposta: (Terminal amino)-metionina-fenilalanina-prolina-lisina-glicina-(terminal carboxila). A sequência de aminoácidos é deduzida usando-se o código genético mostrado no Quadro 12.1. AUG é o códon de iniciação de metionina, seguido do códon de fenilalanina UUU, do códon de prolina CCC, do códon de lisina AAA e do códon de glicina GGG.

3. Se um segmento codificante do filamento molde de um gene (DNA) tem a sequência 3'-TACAAAGGGTTTCCC-5', que

sequência de aminoácidos será produzida se ela for transcrita e traduzida?

Resposta: A sequência de mRNA produzida pela transcrição deste segmento do gene será 5'-AUGUUUCCCAAAGGG-3'. Note que este mRNA tem a mesma sequência de nucleotídeos que aquela discutida no Exercício 2. Assim, irá produzir o mesmo peptídeo quando traduzido: NH₂-Met-Fen-Pro-Lis-Gli-COOH.

4. Que sequência de pares de nucleotídeos em um gene de *Drosophila* codificará a sequência de aminoácidos metionina-triptofano (lendo do terminal amino para o terminal carboxila)?

Resposta: Os códons para metionina e triptofano são AUG e UGG, respectivamente. Assim, a sequência de nucleotídeos no mRNA especificando a sequência de dipeptídeos metionina-triptofano deve ser 5'-AUGUGG-3'. O filamento molde de DNA deve ser complementar e ter polaridade inversa à sequência de mRNA (3'-TACACC-5') e o outro filamento de DNA deve ser complementar ao filamento molde. Portanto, a sequência de pares de bases no gene deve ser:

5'-ATGTGG-3'
3'-TACACC-5'

5. Um gene tipo selvagem contém a sequência de pares de trinucleotídeos:

5'-GAG-3'
3'-CTC-5'

Esta trinca especifica o aminoácido ácido glutâmico. Se o segundo par de bases neste segmento do gene fosse mudado de A:T para T:A, produzindo a seguinte sequência de DNA:

5'-GTG-3'
3'-CAC-5'

ela ainda codificaria o ácido glutâmico?

Resposta: Não; agora ela especificaria o aminoácido valina. O códon para o ácido glutâmico é 5'-GAG-3', que nos diz que o filamento "de baixo" do DNA é o filamento molde. A transcrição do gene tipo selvagem produz a sequência de mRNA 5'-GAG-3', que é um códon de ácido glutâmico. A transcrição do gene alterado produz a sequência de mRNA 5'-GUG-3', que é um códon de valina. De fato, esta é exatamente a mesma mudança de par de nucleotídeos que deu origem à hemoglobina alterada no paciente de Herrick com anemia falciforme, discutida no começo deste capítulo. Veja a Fig. 1.9 para maiores detalhes.

Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. A massa média dos 20 aminoácidos comuns é de cerca de 137 daltons. Avalie o tamanho aproximado de uma molécula de mRNA que codifica um polipeptídeo com uma massa de 65.760 daltons. Suponha que o polipeptídeo contém quantidades iguais de todos os 20 aminoácidos.

Resposta: Baseado nesta suposição, o polipeptídeo conteria cerca de 480 aminoácidos (65.760 daltons/137 daltons por aminoácido). Como cada códon contém três nucleotídeos, a região codificante do mRNA teria que ter 1.440 nucleotídeos de tamanho (480 aminoácidos \times 3 nucleotídeos por aminoácido).

2. O antibiótico estreptomicina mata *E. coli* sensíveis inibindo a ligação de tRNA^{Met} ao sítio P do ribossomo e causando a leitura errada dos códons no mRNA. Em bactérias sensíveis, a estreptomicina é ligada pela proteína S12 à subunidade 30S do ribossomo. A resistência a estreptomicina pode resultar de uma mutação no gene codificante da proteína S12 de modo que a proteína alterada não se liga mais ao antibiótico. Em 1964, Luigi Gorini e Eva Kataja isolaram mutantes de *E. coli* que crescem em meio mínimo suplementado com o aminoácido arginina ou com estreptomicina. Isto é, na ausência de estreptomicina, os mutantes comportam-se como bactérias típicas que precisam de arginina.

Entretanto, na ausência de arginina, elas eram mutantes letais condicionais dependentes de estreptomicina. Isto é, elas cresciam na presença de estreptomicina mas não na ausência de estreptomicina. Explique os resultados obtidos por Gorini e Kataja.

Resposta: Os mutantes letais condicionais dependentes de estreptomicina isolados por Gorini e Kataja continham mutações de sentido trocado em genes que codificavam enzimas biossintéticas de arginina. Se arginina estivesse presente no meio, estas enzimas não eram essenciais. Entretanto, tais enzimas eram necessárias para crescimento na ausência de arginina (um dos 20 aminoácidos necessários para a síntese de proteínas).

A estreptomicina causa leitura errada dos códons de mRNA em bactérias. Essa leitura errada dos códons permitia que os códons que continham as mutações de sentido trocado fossem traduzidos ambigualmente – com a incorporação de aminoácidos errados – quando o antibiótico estava presente. Quando a estreptomicina estava presente na bactéria mutante, um aminoácido ocasionalmente seria inserido (no sítio da mutação), o que resultou em uma enzima ativa que, por sua vez, permitiu que as células crescessem, embora lentamente. Na ausência de estreptomicina, não ocorria erro de leitura, e todos os polipeptídeos mutantes eram inativos.

Questões e Problemas Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 12.1. De modo geral, descreva a organização molecular das proteínas e distinga proteínas do DNA, quimicamente e funcionalmente. Por que a síntese de proteínas é de interesse particular para geneticistas?
- 12.2. Em que local da célula ocorre a síntese de proteínas?
- 12.3. Caracterize os ribossomos em geral quanto a tamanho, localização, função e composição macromolecular.
- 12.4. (a) Onde, nas células de organismos superiores, se originam os ribossomos?
(b) Onde, nas células, os ribossomos são mais ativos na síntese de proteínas?
- 12.5. Identifique três tipos diferentes de RNA que estão envolvidos na tradução e cite as características e funções de cada.
- 12.6. (a) Como o RNA mensageiro está relacionado à formação de polissomo?
(b) Como o rRNA difere do mRNA e do tRNA em especificidade?
(c) Como a molécula de tRNA difere da de DNA e de mRNA em tamanho e arranjo helicoidal?
- 12.7. Esquematize o processo de formação do aminoacil-tRNA.
- 12.8. Que tipos de evidências experimentais foram usados para decifrar o código genético?

- 12.9.** Em que sentido e em que amplitude o código genético é (a) redundante, (b) ordenado e (c) universal?
- 12.10.** O análogo de timina 5-bromouracil é um mutágeno químico que induz substituições de um só par de bases no DNA chamadas transições (substituições de uma purina por outra purina ou uma pirimidina por outra pirimidina). Usando a conhecida natureza do código genético (Quadro 12.1), quais das seguintes substituições de aminoácidos você esperaria que fossem induzidas por 5-bromouracil com a maior frequência:
- (a) Met → Val;
 - (b) Met → Leu;
 - (c) Lis → Tre;
 - (d) Lis → Gln;
 - (e) Pro → Arg; ou
 - (f) Pro → Gln? Por quê?
- 12.11.** Usando a informação dada no Problema 12.10, você esperaria que 5-bromouracil induza uma frequência maior de substituições His → Arg ou His → Pro? Por quê?
- 12.12.** Como a tradução é (a) iniciada e (b) terminada?
- 12.13.** Se a massa molecular média de um aminoácido é supostamente de 100 dáltons, quantos nucleotídeos estarão presentes em uma sequência codificante de mRNA que especifica um único polipeptídeo com uma massa molecular de 27.000 dáltons?
- 12.14.** Qual o significado da hipótese da oscilação?
- 12.15.** As bases A, G, U, C, I (inosina) ocorrem nas posições 5' de anticódons nos tRNA.
- (a) Que base pode parear-se com três bases diferentes nas posições 3' dos códons no mRNA?
 - (b) Qual é o número mínimo de tRNA necessário para reconhecer todos os códons de aminoácidos especificados por códons com redundância completa?
- 12.16.** Qual o número mínimo de tRNA necessário para reconhecer os seis códons que especificam o aminoácido leucina?
- 12.17.** Suponha que, no ano 2025, a primeira expedição de humanos a Marte descubra várias formas de vida marciana que se nutrem de fontes hidrotérmicas que existem abaixo da superfície do planeta. Várias equipes de biólogos moleculares extraem proteínas e ácidos nucleicos desses organismos e fazem importantes descobertas. Sua primeira descoberta é que as proteínas nas formas de vida marcianas contêm apenas 14 aminoácidos diferentes em vez dos 20 presentes nas formas de vida de nosso planeta. Sua segunda descoberta é que o DNA e o RNA nesses organismos têm apenas dois nucleotídeos diferentes em lugar dos quatro nucleotídeos presentes em organismos vivos na Terra. (a) Supondo que a transcrição e a tradução funcionam similarmente nos marcianos e nos terráqueos, qual o número mínimo de nucleotídeos que devem estar presentes nos códons marcianos para especificar todos os seus aminoácidos? (b) Supondo que o código marciano proposto acima tenha sinais de início e de fim, você esperaria que o código genético marciano seja redundante como o nosso?
- 12.18.** (a) Por que o código genético é triplo em vez de unitário ou duplo?
- (b) Quantos aminoácidos diferentes são especificados pelo código genético?
 - (c) Quantas sequências diferentes de aminoácidos são possíveis em um polipeptídeo com 146 aminoácidos de tamanho?
- 12.19.** Quais as diferenças básicas entre tradução em procariontes e em eucariontes?
- 12.20.** Qual a função de cada um dos seguintes componentes do aparelho de síntese de proteínas:
- (a) aminoacil-tRNA-sintetase,
 - (b) fator I de liberação,
 - (c) peptidiltransferase,
 - (d) fatores de iniciação,
 - (e) fator de alongamento G?
- 12.21.** Um gene de *E. coli* foi isolado tendo sido demonstrado possuir 68 nm de tamanho. Qual o número máximo de aminoácidos que este gene pode codificar?
- 12.22.** (a) Qual a diferença entre uma mutação sem sentido e uma mutação de sentido trocado?
- (b) As mutações sem sentido ou de sentido trocado são mais frequentes em organismos vivos?
 - (c) Por quê?
- 12.23.** A cadeia de α -globina humana tem 141 aminoácidos de tamanho. Quantos nucleotídeos no mRNA são necessários para codificar esta cadeia?
- 12.24.** Quais as funções dos sítios de ligação *A*, *P* e *E* do aminoacil-tRNA no ribossomo?
- 12.25.** (a) De que modos a ordem no código genético minimiza a letalidade mutacional?
- (b) Por que as mudanças de par de bases que causam a substituição de uma leucina por uma valina no produto gênico polipeptídico raramente produzem um fenótipo mutante?
- 12.26.** (a) Qual a função da sequência Shine-Dalgarno nos mRNA procarióticos?
- (b) Que efeito a deleção da sequência Shine-Dalgarno de um mRNA tem em sua tradução?
- 12.27.** (a) De que modos ribossomos e spliceossomos são similares?
- (b) De que modos são diferentes?
- 12.28.** Se você fosse (1) purificar um tRNA de cisteína e carregá-lo com cisteína marcada (isto é, ativá-lo ligando cisteína marcada com ^3H), (2) usar níquel de Raney (um pó de níquel altamente redutor) para converter a cisteína em alanina ainda ligada ao tRNA específico de cisteína e (3) colocar o tRNA carregado com alanina em um sistema de síntese de proteínas *in vitro* ativado com moldes de poli(UG) que normalmente estimulam a incorporação de cisteína mas não alanina nas cadeias polipeptídicas (isto é, quando você usa

tRNA normais carregados com alanina e tRNA carregado com cisteína), que resultado você espera?

- 12.29.** Uma sequência parcial de nucleotídeos (5' subterminal) de um mRNA procariótico é a seguinte:

5'...AGGAGGCUCGAACAUGUCAUAUGCUUGUUC
CAAUCGUUAGCUGCGCAGGACCGUCCCGGA...3'

Quando este mRNA é traduzido, que sequência de aminoácidos será especificada por esta parte do mRNA?

- 12.30.** O terminal 5' de um mRNA humano tem a seguinte sequência:

5'*cap*-GAAGAGACAAGGTCAUGGCCAUAUGCUUGU-
UCCAAUCGUUAGCUGCGCAGGAUCGCCCUGGG...3'

Quando este mRNA é traduzido, que sequência de aminoácidos será especificada por esta parte do mRNA?

- 12.31.** Alan Garen estudou amplamente uma determinada mutação sem sentido (término de cadeia) no gene de fosfatase alcalina de *E. coli*. Esta mutação resultou no término da cadeia polipeptídica de fosfatase alcalina em uma posição onde o aminoácido triptofano ocorreu no polipeptídeo tipo selvagem. Garen induziu revertentes (neste caso, mutações que alteram o mesmo códon) deste mutante com mutágenos químicos que induziram substituições de um único par de bases e sequenciaram os polipeptídeos nos revertentes. Foram encontrados sete tipos diferentes de revertentes, cada um com um aminoácido diferente na posição do triptofano do polipeptídeo tipo selvagem (posição de término do fragmento polipeptídico mutante). Os aminoácidos presentes nesta posição nos vários revertentes incluíam tripto-

fano, serina, tirosina, leucina, ácido glutâmico, glutamina e lisina. A mutação sem sentido estudada por Garen era uma mutação *âmbar* (UAG), *ocre* (UAA) ou *opala* (UGA)? Explique a base de sua dedução.

- 12.32.** A seguinte sequência de DNA ocorre em uma bactéria (a sequência promotora está situada à esquerda mas não é mostrada).

↓
5'-CAATCATGGACTGCCATGCTTCATATGAATAGTTGACAT-3'
3'-GTTAGTACCTGACGGTACGAAGTATACTTATCAACTGTA-5'

- (a) Qual a sequência de ribonucleotídeos da molécula de mRNA que é transcrita do filamento molde deste trecho de DNA? Suponha que estão presentes tanto o códon de início quanto o de término.
- (b) Qual a sequência de aminoácidos do polipeptídeo codificado por este mRNA?
- (c) Se o nucleotídeo indicado pela seta sofre uma mutação que faz com que este par de bases C:G seja deletado, qual será o polipeptídeo codificado pelo gene mutante?
- 12.33.** A seguinte sequência de DNA ocorre no filamento não-molde de um gene estrutural em uma bactéria (a sequência promotora está situada a esquerda mas não é mostrada):
- ↓
5'-GAATGTCAGAACTGCCATGCTTCATATGAATAGACCTCTAG-3'
- (a) Qual a sequência de ribonucleotídeos da molécula de mRNA que é transcrita deste trecho de DNA?
- (b) Qual a sequência de aminoácidos do polipeptídeo codificado por este mRNA?
- (c) Se o nucleotídeo indicado pela seta sofre uma mutação que muda T para A, qual será a sequência de aminoácidos resultante após a transcrição e a tradução?

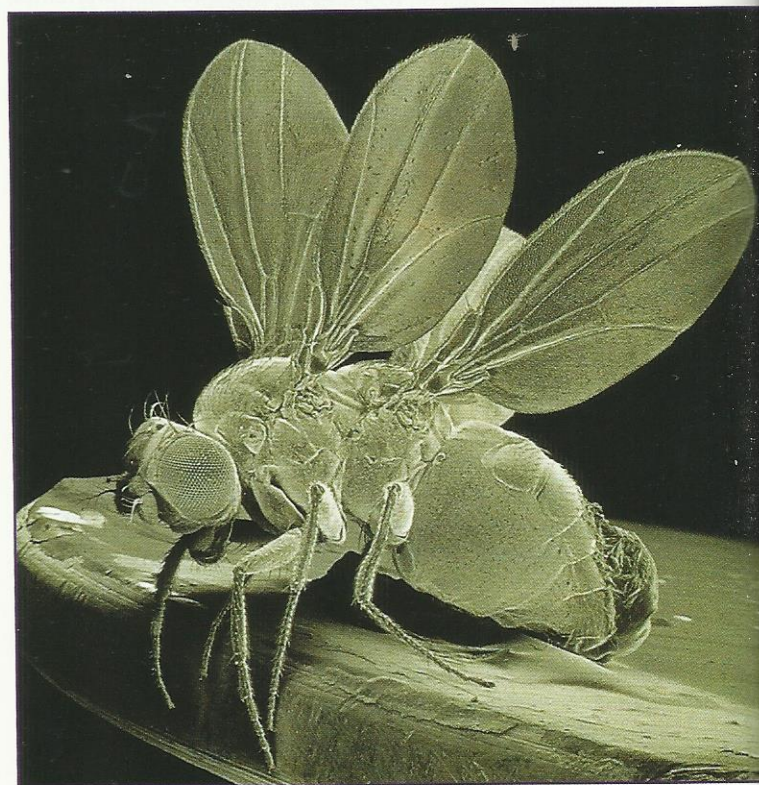
Mutação, Reparo do DNA e Recombinação

Tópicos

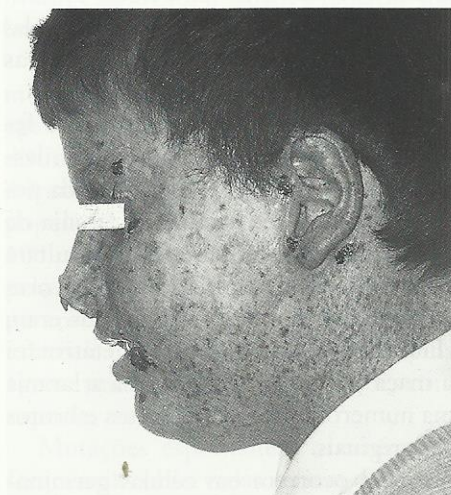
- **Mutação: Fonte da Variabilidade Genética Necessária para Evolução**
- **Mutação: Características Básicas do Processo**
- **Mutação: Efeitos Fenotípicos**
- **A Base Molecular da Mutação**
- **Triagem de Substâncias Químicas Quanto a Mutagenicidade: O Teste de Ames**
- **Mecanismos de Reparo do DNA**
- **Doenças Humanas Herdadas com Defeitos no Reparo do DNA**
- **Mecanismos de Recombinação do DNA**

Xeroderma Pigmentoso: Reparo Defeituoso do DNA Danificado em Humanos

O sol brilhava intensamente em dia da metade do verão, um dia perfeito para a maioria das crianças passar na praia. Todos os amigos de Nathan vestiam roupas de banho ou *shorts*. Preparando-se para se juntar a seus amigos, ele vestiu calças compridas e camisas de mangas longas. Ele colocou um chapéu de aba larga e passou uma grossa camada de protetor solar em suas mãos e na face. Enquanto os amigos de Nathan adoram brincar



Micrografia eletrônica de varredura, feita por Edward Lewis, de uma mosca-das-frutas com quatro asas produzida pela combinação de três mutações no locus *bitorax*.



ao ar livre, ele vive em constante temor dos efeitos da luz solar. Ele nasceu com o distúrbio herdado xeroderma pigmentoso, uma característica autossômica recessiva que afeta cerca de uma em cada 250.000 crianças. As células da pele de Nathan são extremamente sensíveis à radiação ultravioleta, os raios de alta energia do sol. A luz ultravioleta causa mudanças químicas no DNA das células da pele de Nathan, mudanças que resultam não só em intensas sardas como também em câncer de pele (Fig. 13.1).

Os amigos de Nathan não se preocupam em brincar ao sol; queimaduras são a sua única preocupação maior. Suas células da pele contêm enzimas que corrigem as mudanças no DNA resultantes de exposição a luz ultravioleta. Entretanto, as células da pele de Nathan não têm uma das enzimas necessárias para o reparo das alterações induzidas pela luz ultravioleta na estrutura do DNA. O xeroderma pigmentoso pode resultar de defeitos herdados em qualquer um dentre nove genes humanos diferentes. Além disso, sabe-se que outros distúrbios herdados resultam da falha de reparo do DNA danificado por outros agentes físicos e químicos. As consequências ameaçadoras à vida destes defeitos herdados nas enzimas de reparo do DNA destacam dramaticamente sua importância.

Fig. 13.1 ■ Efeitos fenotípicos da doença hereditária xeroderma pigmentoso. Pessoas com esta doença maligna desenvolvem amplos tumores de pele após exposição ao sol.

Considerando o papel fundamental que o DNA tem no crescimento e no desenvolvimento de organismos vivos, a evolução de mecanismos para proteger sua integridade parece inevitável. De fato, como discutimos neste capítulo, as células vivas contêm várias enzimas que constantemente percorrem o DNA à procura de pares de nucleotídeos incorretos ou danificados. Quando detectados, esses defeitos são corrigidos por um pequeno exército de **enzimas de reparo do DNA**, cada uma tendo se desenvolvido para combater um tipo particular de dano. Neste capítulo, examinaremos os tipos de mudanças que ocorrem no DNA, os processos pelos quais essas alterações são corrigidas e os processos correlatos de recombinação entre moléculas homólogas de DNA.

MUTAÇÃO: FONTE DA VARIABILIDADE GENÉTICA NECESSÁRIA PARA EVOLUÇÃO

As mutações, mudanças herdadas no material genético, fornecem uma grande variação genética que permite a evolução de organismos.

Sabemos pelos capítulos anteriores que a herança é baseada em genes que são transmitidos de genitores para a prole durante a reprodução e que os genes estocam informação genética codificada nas sequências de pares de nucleotí-

deos no DNA ou nucleotídeos no RNA. Examinamos como esta informação genética é duplicada com precisão durante a replicação semiconservativa do DNA. Foi demonstrado que essa replicação precisa depende em parte das atividades de revisão inerentes às DNA-polimerases que catalisam a síntese de DNA. Assim, desenvolveram-se mecanismos para facilitar a transmissão fiel da informação genética de célula a célula e finalmente de geração a geração. Entretanto, ocorrem erros no material genético. Tais mudanças herdáveis no material genético são chamadas de mutações.

O termo **mutação** refere-se tanto (1) a mudança no material genético quanto (2) ao processo pelo qual ocorre a mudança. Um organismo que apresenta um fenótipo novo resultante de uma mutação é chamado **mutante**. Usada no sentido histórico mais amplo, a mutação refere-se a qualquer mudança, súbita e herdável, no genótipo de uma célula ou um organismo. Entretanto, as mudanças no genótipo, e portanto no fenótipo, de um organismo que resultam de eventos de recombinação que produzem novas combinações de variação genética já existente devem ser cuidadosamente distinguidas de mudanças causadas por mutações novas. Ambos os eventos às vezes dão origem a fenótipos novos em frequências muito baixas. Alterações mutacionais no genótipo de um organismo incluem mudanças no número e na estrutura dos cromossomos (Cap. 6), bem como mudanças nas estruturas de genes individuais. Mutações que envolvem alterações em sítios específicos em um gene são denominadas **mutações de ponto**. Elas incluem a substituição de um par de bases por outro ou a

inserção ou deleção de um ou alguns pares de nucleotídeos em um local específico de um gene. Hoje, o termo *mutação* geralmente é usado em um sentido restrito para se referir apenas a mutações de ponto. Neste capítulo, exploraremos o processo de mutação conforme definido no sentido restrito.

Mutação é a principal fonte de toda variação genética. Ela fornece a matéria-prima para a evolução. Os mecanismos de recombinação rearranjam a variabilidade genética em novas combinações, e a seleção natural ou artificial preserva as combinações mais bem adaptadas às condições existentes no ambiente ou desejadas pelos que fazem cruzamentos de animais ou plantas. Sem mutações, todos os genes existiriam em apenas uma forma. Alelos não existiriam, e a análise genética clássica não seria possível. Mais importante, populações de organismos não seriam capazes de evoluir e se adaptar a mudanças ambientais. Algum nível de mutação é essencial para dar nova variabilidade genética e permitir que os organismos se adaptem a novos ambientes. Ao mesmo tempo, se mutações ocorressem com muita frequência, elas perturbariam a fiel transmissão da informação genética de geração a geração. Além disso, a maioria das mutações com efeitos fenotípicos facilmente detectados são deletérias aos organismos nos quais elas ocorrem. Como seria esperado, a taxa de mutação é influenciada por fatores genéticos, e desenvolveram-se mecanismos que regulam o nível de mutação que ocorre em várias condições ambientais.

■ PONTO IMPORTANTE

- Mutações são alterações herdáveis no material genético que fornecem a matéria-prima para evolução.

MUTAÇÃO: CARACTERÍSTICAS BÁSICAS DO PROCESSO

Mutações ocorrem em todos os organismos, de vírus a humanos. Elas podem ocorrer espontaneamente ou ser induzidas por agentes mutagênicos. A mutação é geralmente um processo aleatório, não adaptativo.

Mutações ocorrem em todos os genes de todos os organismos vivos. Essas mutações fornecem nova variabilidade genética que permite aos organismos adaptar-se a mudanças ambientais. Assim, as mutações foram, e continuam a ser, essenciais ao processo evolutivo. Antes de discutirmos os efeitos fenotípicos específicos das mutações e os mecanismos pelos quais ocorrem os vários tipos de mutação, consideraremos algumas das características básicas deste processo importante.

Mutação: Somática ou Germinativa

Uma mutação pode ocorrer em qualquer célula em qualquer estágio do desenvolvimento de um organismo multi-

celular. Os efeitos imediatos da mutação e sua capacidade de produzir uma mudança fenotípica são determinados por sua dominância, pelo tipo de célula na qual ocorre e pela época em que ocorre durante o ciclo de vida do organismo. Em animais superiores, as células da linhagem germinativa que dão origem aos gametas separam-se de outras linhagens celulares no início do desenvolvimento (Cap. 2). Todas as células das linhagens não germinativas são células somáticas. **Mutações germinativas** são as que ocorrem em células da linhagem germinativa, enquanto **mutações somáticas** ocorrem em células somáticas.

Se uma mutação ocorre em uma célula somática, o fenótipo mutante resultante se manifestará apenas nos descendentes desta célula. A mutação não será transmitida pelos gametas para a prole. A maçã Delicious e a laranja de umbigo são exemplos de fenótipos mutantes que resultam de mutações que ocorrem em células somáticas. As árvores dos frutos nas quais ocorreram as mutações originais eram mosaicos somáticos. Felizmente, a propagação vegetativa foi possível tanto para a maçã Delicious quanto para a laranja de umbigo, e hoje uma numerosa prole de enxertos e brotos perpetuou as mutações originais.

Se mutações dominantes ocorrem em células germinativas, seus efeitos podem ser expressos imediatamente na prole. Se as mutações são recessivas, seus efeitos em geral são obscurecidos em diplóides. Mutações germinativas podem ocorrer em qualquer estágio no ciclo reprodutivo do organismo. Se a mutação surge em um gameta, apenas um único membro da prole provavelmente terá o gene mutante. Se uma mutação ocorre em uma célula primordial da linhagem germinativa do testículo ou ovário, vários gametas podem receber o gene mutante, acentuando seu potencial de perpetuação. Assim, a dominância de um alelo mutante e o estágio do ciclo reprodutivo no qual ocorreu uma mutação são os principais fatores para determinar a probabilidade de que o alelo mutante se manifeste em um organismo.

A mais antiga mutação germinativa dominante registrada em animais domésticos foi observada por Seth Wright em 1791 em sua fazenda em Charles River em Dover, Massachusetts. Entre seu rebanho de ovelhas, Wright notou uma peculiar com pernas incomumente curtas. Ele logo imaginou que seria vantajoso ter um rebanho inteiro destas ovelhas de pernas curtas, que não poderiam pular as cercas baixas da vizinhança da Nova Inglaterra. Wright usou as ovelhas de pernas curtas para cruzarem na estação seguinte. Duas de suas ovelhas tinham pernas curtas. As ovelhas de pernas curtas foram entrecruzadas, e foi desenvolvida uma linhagem na qual a nova característica foi expressa em todos os animais.

Mutação: Espontânea ou Induzida

Quando ocorre uma nova mutação, tal como a que produziu ovelhas de Wright com pernas curtas, ela é causada por algum agente no ambiente ou resulta de um processo inerente em organismos vivos? **Mutações espontâneas** são aque-

las que ocorrem sem uma causa conhecida. Elas podem ser verdadeiramente espontâneas, resultando de um nível baixo de erros metabólicos inerentes, ou podem ser causadas por agentes desconhecidos presentes no ambiente. **Mutações induzidas** são aquelas resultantes da exposição de organismos a agentes físicos e químicos que causam mudanças no DNA (ou RNA em alguns vírus). Tais agentes são chamados **mutágenos**. Eles incluem irradiação ionizante, luz ultravioleta e uma ampla variedade de substâncias.

Operacionalmente, é impossível provar que determinada mutação ocorreu espontaneamente ou foi induzida por um agente mutagênico. Os geneticistas devem restringir tais distinções em nível populacional. Se uma taxa de mutação é aumentada uma centena de vezes por tratamento de uma população com um mutágeno, uma média de 99 de cada 100 mutações presentes na população terão sido induzidas pelo mutágeno. Os pesquisadores podem então fazer comparações válidas entre mutações espontâneas e induzidas estatisticamente comparando as populações expostas a um agente mutagênico com populações controles que não foram expostas ao mutágeno.

Mutações espontâneas ocorrem raramente, embora as frequências observadas variem de gene para gene e de organismo para organismo. As medidas das frequências de mutações espontâneas para vários genes de fagos e bactérias variam de cerca de 10^{-8} a 10^{-10} mutações detectáveis por par de nucleotídeos por geração. Para eucariontes, as estimativas das taxas de mutação variam de cerca de 10^{-7} a 10^{-9} mutações detectáveis por par de nucleotídeos por geração (considerando apenas os genes para os quais amplos dados estão disponíveis). Comparando taxas de mutação por nucleotídeo com taxas de mutação por gene, a região codificante do gene médio geralmente é tida como sendo de 1.000 pares de nucleotídeos de tamanho. Assim, a taxa de mutação por gene varia de cerca de 10^{-4} a 10^{-7} por geração.

O tratamento com agentes mutagênicos pode aumentar as frequências de mutação em ordens de magnitude. A frequência de mutação por gene em bactérias e vírus pode ser aumentada em mais de 1 por cento pelo tratamento com mutágenos químicos potentes. Isto é, mais de 1 por cento dos genes dos organismos tratados conterão uma mutação, ou, dito de outro modo, mais de 1 por cento dos fagos ou bactérias na população terão uma mutação em determinado gene.

Mutação: Geralmente um Processo Aleatório, Não Adaptativo

Os ratos em muitas cidades não são mais afetados por anticoagulantes que tradicionalmente tinham sido usados como venenos de roedores. Muitas populações de baratas são insensíveis a clordano, o veneno usado para controlá-las nos anos 1950. As populações de moscas domésticas apresentam níveis altos de resistência a muitos inseticidas. Cada vez mais microrganismos patogênicos estão se tornando resistentes

a antibióticos desenvolvidos para controlá-los. A introdução desses pesticidas e antibióticos pelos humanos produziu novos ambientes para tais organismos. Mutações que produzem resistência a esses pesticidas e antibióticos ocorrem; os organismos sensíveis foram mortos; e os mutantes multiplicaram-se para produzir novas populações resistentes. Muitos desses casos de evolução por mutação e seleção natural estão bem documentados.

Estes exemplos levantam uma dúvida básica sobre a natureza da mutação. A mutação é um fenômeno puramente aleatório no qual o estresse ambiental apenas preserva mutações preexistentes? Ou a mutação é dirigida pelo estresse ambiental? Por exemplo, se você cortar a cauda de camundongos por várias gerações, você produzirá uma linhagem de camundongos sem cauda? A despeito da crença de Jean Lamarck e Trofim Lysenko, que acreditavam na herança de "características adquiridas" – características impostas a organismos por fatores ambientais –, a resposta é não. Os camundongos continuarão a nascer com caudas.

Hoje em dia, é difícil compreender como Lysenko apresentou sua crença no *lamarckismo*, a herança de caracteres adquiridos, aos poderosos da União Soviética de 1937 a 1964. Entretanto, desmentir o *lamarckismo* não foi uma tarefa fácil, especialmente no caso de microrganismos, em que mesmo pequenas culturas em geral contêm bilhões de organismos.

Como um exemplo, consideremos uma população de bactérias tais como *E. coli* crescendo em um ambiente livre de estreptomicina. Quando expostas a estreptomicina, a maioria das bactérias serão mortas pelo antibiótico. Entretanto, se a população for grande o suficiente, ela logo dará origem a uma cultura resistente a estreptomicina na qual todas as células são resistentes ao antibiótico. A estreptomicina simplesmente seleciona mutantes raros de ocorrência aleatória preexistentes na população, ou todas as células têm alguma probabilidade baixa de desenvolver resistência em resposta à presença de estreptomicina? Como os geneticistas distinguem estas duas possibilidades? Resistência a estreptomicina só pode ser detectada tratando-se a cultura com o antibiótico. Como, então, um geneticista pode determinar se bactérias resistentes estão presentes antes da exposição a estreptomicina ou são induzidas pela presença do antibiótico?

Em 1952, Joshua e Esther Lederberg desenvolveram uma nova técnica importante chamada **plaqueamento réplica**. Esta técnica permitiu-lhes demonstrar a presença de mutantes resistentes a antibióticos em culturas bacterianas antes de exposição ao antibiótico (Fig. 13.2). Os Lederberg primeiro diluíram as culturas bacterianas, espalharam as bactérias na superfície de um meio nutriente de ágar semi-sólido em placas de Petri e incubaram as placas até que cada bactéria tivesse produzido uma colônia visível na superfície do ágar. Em seguida, eles inverteram cada placa e a pressionaram sobre um veludo estéril colocado em bloco de madeira. Algumas das células de cada colônia grudaram no veludo. Eles então pressionaram uma placa de nutriente estéril de meio de ágar contendo estreptomicina sobre o veludo. Eles repetiram este procedimento de plaqueamento réplica com

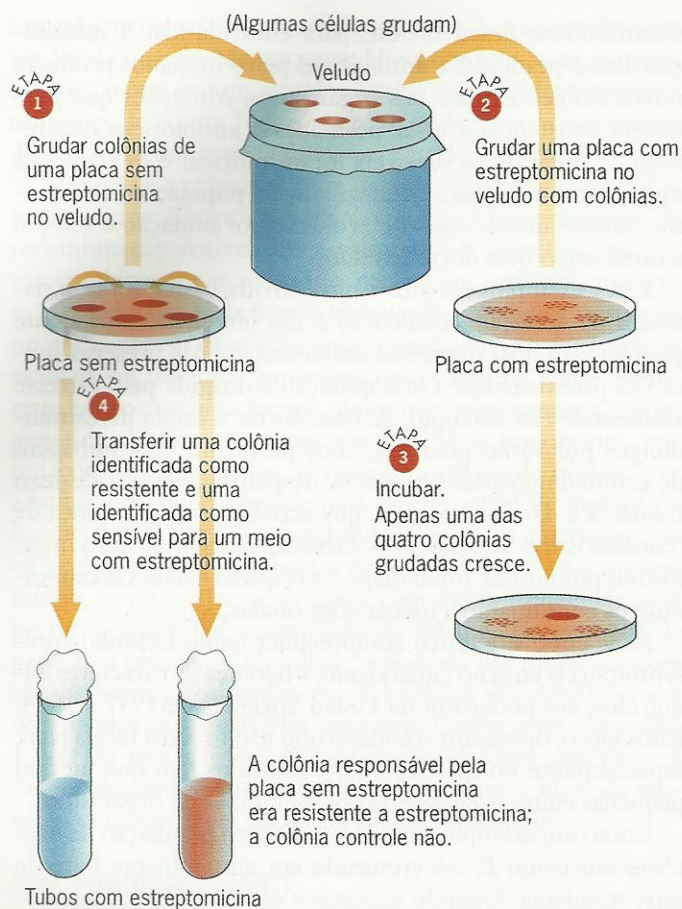


Fig. 13.2 ■ Joshua e Esther Lederberg usaram o plaqueamento réplica para demonstrar a natureza aleatória ou não direcional da mutação. Por simplificação, apenas quatro colônias são mostradas em cada placa e apenas duas são testadas quanto à resistência a estreptomicina na etapa 4. De fato, cada placa conterá cerca de 200 colônias, e muitas placas seriam usadas para encontrar um número adequado de colônias mutantes.

muitas placas, cada uma contendo cerca de 200 colônias bacterianas. Após incubarem as placas seletivas (aquelas contendo estreptomicina) de um dia para outro, raras colônias resistentes a estreptomicina tinham se formado.

Os Lederberg subsequentemente testaram as colônias em placas não seletivas (as que não continham estreptomicina) quanto a sua capacidade de crescer em meio contendo estreptomicina. Seus resultados foram definitivos. As colônias que cresceram em placas réplicas seletivas quase sempre continham células resistentes a estreptomicina, enquanto as que não cresciam no meio seletivo raramente continham qualquer célula resistente (Fig. 13.2).

Se uma mutação que torna uma bactéria resistente a estreptomicina ocorre em um estágio inicial no crescimento de uma colônia, a célula resistente irá dividir-se e produzirá duas, então quatro, então oito e no final um grande número de bactérias resistentes. Assim, se uma mutação estiver ocorrendo aleatoriamente, processo não adaptativo, muitas das

colônias que são formadas em placas não seletivas conterão mais que uma bactéria resistente a antibióticos e darão origem a culturas resistentes quando testadas quanto a crescimento em meio seletivo. Entretanto, se uma mutação é adaptativa e as mutações para resistência a estreptomicina só ocorrem após exposição ao antibiótico, então as colônias nas placas não seletivas que dão origem a colônias resistentes nas placas seletivas após plaqueamento réplica provavelmente não conterão mais células resistentes a estreptomicina do que as outras colônias nas placas não seletivas.

Assim, usando sua técnica de plaqueamento réplica, os Lederberg demonstraram a existência de mutantes resistentes a estreptomicina em uma população de bactérias antes de sua exposição ao antibiótico. Seus resultados, juntamente com os de muitos outros experimentos, mostraram que o estresse ambiental não dirige ou causa mudanças genéticas como pensava Lysenko. Ele simplesmente seleciona raras mutações preexistentes que resultam em fenótipos mais bem adaptados ao novo ambiente.

Mutagênese Adaptativa, ou de Fase Estacionária, em Bactérias

Uma **mutação adaptativa** fornece uma vantagem seletiva ao organismo mutante quando cultivado no ambiente em que se originou. A formação de mutações adaptativas em bactérias é talvez mais apropriadamente chamada **mutagênese de fase estacionária** porque ocorre quando populações de bactérias param de crescer – entram em uma fase estacionária – devido a inanição ou algum outro estresse ambiental (Fig. 13.3). Bactérias na fase estacionária ou param de se dividir ou morrem na mesma taxa em que novas células são

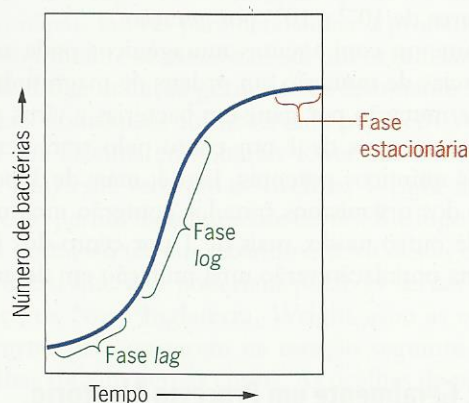


Fig. 13.3 ■ Uma curva de crescimento bacteriano. A transferência de bactérias sem crescimento para um novo meio nutriente resulta em três estágios sucessivos de crescimento: fase lag, fase log e fase estacionária. Durante a fase lag, as bactérias ajustam seu metabolismo para facilitar o crescimento no novo meio. Uma vez feito este ajuste, começa o crescimento exponencial que continua durante a fase log. Quando um ou mais nutrientes necessários no meio se tornam limitantes de crescimento, este diminui e acaba parando na fase estacionária.

produzidas. As mutações adaptativas nestas células de fase estacionária são produzidas juntamente com outras mutações que ocorrem aleatoriamente (mutações não adaptativas); isto é, resultam de um aumento induzido por estresse na taxa de mutação. Essa mutagênese aumentada, por sua vez, é causada pela indução de processos de reparo do DNA propensos a erro (veja a seção Mecanismos de Reparo do DNA, mais adiante neste capítulo).

Devemos destacar que “adaptativo” não significa que um tipo específico de mutação que melhora o crescimento sob um determinado estresse ambiental foi preferencialmente induzido por tal estresse ambiental. Esta última seria uma “mutação dirigida”, para a qual não há evidência válida. Em vez disso, “adaptativa” significa que algumas mutações que dão uma vantagem seletiva à bactéria ocorreram em resposta ao estresse, juntamente com muitas outras mutações não adaptativas. De fato, há evidência de mutagênese adaptativa, ou de fase estacionária, em várias espécies de bactérias. Entretanto, a maioria das pesquisas deste fenômeno foram feitas com *E. coli* e *Bacillus subtilis*. Mesmo nestas duas espécies, a mutagênese de fase estacionária não é universal; ela ocorre em algumas linhagens e não em outras. Um estudo examinou 787 linhagens de *E. coli* que foram isoladas de uma ampla faixa de ambientes naturais em todo o mundo. Quando essas linhagens foram testadas quanto a mutagênese de fase estacionária durante sete dias de inanição, 80% apresentaram algum nível de mutagênese de fase estacionária. Entretanto, a frequência de mutações induzidas variou muito de linhagem para linhagem.

Quando bactérias tais como *E. coli* entram na fase estacionária devido a inanição ou outro estresse ambiental, é induzida uma via de reparo do DNA propensa a erro chamada resposta SOS (veja Mecanismos de Reparo do DNA). Durante a resposta SOS, um grande número de genes que codificam proteínas envolvidas no metabolismo do DNA – replicação, recombinação e reparo – são ativados. Algumas das proteínas induzidas estão envolvidas no reparo do DNA danificado por recombinação. Outras são DNA-polimerases propensas a erro (IV e V em *E. coli*) que replicam através de segmentos danificados do DNA e, desta forma, produzem mutações. A resposta SOS é claramente responsável por algumas mutações de fase estacionária. Entretanto, estudos recentes mostraram que a mutagênese de fase estacionária ocorre em linhagens bacterianas que não têm componentes essenciais do sistema de resposta SOS. Outros estudos forneceram evidências de várias vias de mutagênese de fase estacionária.

No momento, podemos concluir que a mutagênese adaptativa, ou de fase estacionária, é generalizada em bactérias, e contribui para sua capacidade de se adaptar a diversos ambientes. Entretanto, os mecanismos moleculares pelos quais ocorre a mutagênese de fase estacionária precisam de outros estudos, como a questão sobre se o fenômeno ocorre em outras espécies. Embora existam evidências de mutagênese de fase estacionária em leveduras, e DNA-polimerases propensas a erro estejam presentes em eucariontes, a possibilidade

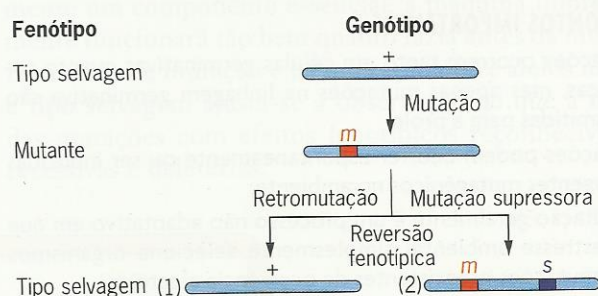
de que processos similares ocorram em eucariontes permanece ainda amplamente não estudada.

Mutação: Um Processo Reversível

Como já discutimos, uma mutação em um gene tipo selvagem pode produzir um alelo mutante que resulta em um fenótipo anormal. Entretanto, o alelo mutante também pode sofrer uma mutação reversa para uma forma que restaura o fenótipo tipo selvagem. Isto é, a mutação é um processo reversível.

A mutação de um gene tipo selvagem para uma forma que resulta em um fenótipo mutante é referida como **mutação direta**. Entretanto, às vezes a designação dos fenótipos tipo selvagem e mutante é bem arbitrária. Eles podem simplesmente representar dois fenótipos diferentes mas normais. Por exemplo, os geneticistas consideram os alelos para as cores de olho castanho e azul em humanos como sendo ambos tipo selvagem. Entretanto, em uma população composta quase totalmente de pessoas com olhos castanhos, o alelo para olhos azuis poderia ser tido como um alelo mutante.

Quando uma segunda mutação restaura o fenótipo original perdido por causa de uma mutação anterior, o processo



Retrocruzamento com tipo selvagem:

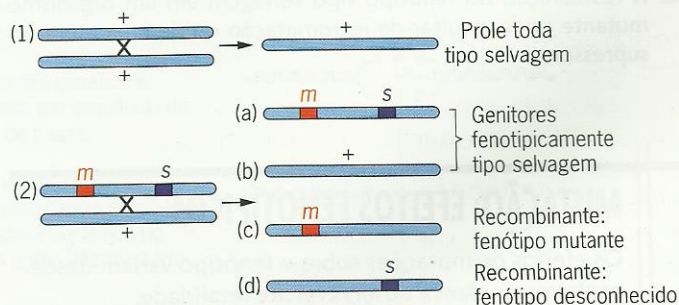


Fig. 13.4 ■ Restauração do fenótipo original tipo selvagem em um organismo pode ocorrer por (1) mutação reversa ou (2) mutação supressora (mostrada no mesmo cromossomo por simplificação). Alguns mutantes podem reverter ao fenótipo tipo selvagem por ambos os mecanismos. Os revertentes dos dois tipos podem ser distinguidos por retrocruzamentos com o tipo selvagem original. Se ocorreu uma mutação reversa, toda a prole do retrocruzamento será tipo selvagem. Se uma mutação supressora é responsável, parte da prole do retrocruzamento terá o fenótipo mutante (2c).

é chamado de **mutação reversa** ou **mutação de reversão**. A reversão pode ocorrer de dois modos diferentes: (1) por **mutação reversa**, uma segunda mutação no mesmo ponto no gene que a mutação original, restaurando a sequência nucleotídica tipo selvagem, ou (2) pela ocorrência de uma **mutação supressora**, uma segunda mutação em um local diferente no genoma, que compensa os efeitos da primeira mutação (Fig. 13.4). A mutação reversa restaura a sequência original tipo selvagem de nucleotídeos do gene, enquanto a mutação supressora não. Mutações supressoras podem ocorrer em pontos distintos no mesmo gene que a mutação original ou em genes diferentes, até mesmo em cromossomos diferentes.

Algumas mutações reverterem primariamente por retro-mutação, enquanto outras o fazem quase exclusivamente pela ocorrência de mutações supressoras. Assim, em estudos genéticos, os pesquisadores em geral devem distinguir entre estas duas possibilidades retrocruzando o revertente fenotípico com o organismo original tipo selvagem. Se o fenótipo tipo selvagem é restaurado por uma mutação supressora, a mutação original ainda estará presente e pode ser separada da mutação supressora por recombinação (Fig. 13.4). Se o fenótipo tipo selvagem é restaurado por retromutação, toda a prole do retrocruzamento será tipo selvagem.

PONTOS IMPORTANTES

- Mutações ocorrem tanto em células germinativas quanto somáticas, mas apenas mutações na linhagem germinativa são transmitidas para a prole.
- Mutações podem ocorrer espontaneamente ou ser induzidas por agentes mutagênicos no ambiente.
- A mutação geralmente é um processo não adaptativo em que um estresse ambiental simplesmente seleciona organismos com mutações preexistentes de ocorrência aleatória.
- A mutagênese adaptativa, ou de fase estacionária, ocorre em bactérias que foram expostas a um estresse ambiental como inanição.
- A restauração do fenótipo tipo selvagem em um organismo mutante pode resultar de retromutação ou de uma mutação supressora.

MUTAÇÃO: EFEITOS FENOTÍPICOS

Os efeitos de mutações sobre o fenótipo variam desde nenhuma mudança observável até letalidade.

Os efeitos de mutações sobre o fenótipo variam de alterações tão pequenas que só podem ser detectadas por técnicas genéticas ou bioquímicas especiais a grandes modificações morfológicas e a alterações letais. Um gene é uma sequência de pares de nucleotídeos que geralmente codificam um polipeptídeo específico. Qualquer mutação que ocorra dentro de determinado gene produzirá um novo alelo desse gene. Os genes contendo mutações sem efeito sobre o

fenótipo, ou pequenos efeitos que só podem ser reconhecidos por técnicas especiais, são chamados **isoalelos**. Outras mutações produzem **alelos nulos** que resultam na falta de produto gênico ou produtos gênicos totalmente não funcionais. Se mutações deste último tipo ocorrem em genes que são necessários para o crescimento do organismo, os indivíduos que são homozigotos para a mutação não sobreviverão. Tais mutações são chamadas de **letais recessivos**.

As mutações podem ser recessivas ou dominantes. Em organismos monoplóides tais como vírus e bactérias, tanto mutações recessivas quanto dominantes podem ser reconhecidas por seu efeito sobre o fenótipo do organismo no qual elas ocorrem. Em organismos diplóides tais como a mosca-das-frutas e seres humanos, as mutações recessivas irão alterar o fenótipo apenas quando presentes na condição homozigota. Assim, em diplóides, a maioria das mutações recessivas não serão reconhecidas na época de sua ocorrência, pois estarão presentes no estado heterozigoto. Mutações recessivas ligadas ao X são uma exceção; elas serão expressas no estado hemizigoto no sexo heterogamético (por exemplo, indivíduos do sexo masculino em seres humanos e na mosca-das-frutas e do sexo feminino em aves). Mutações letais recessivas ligadas ao X irão alterar a proporção dos sexos na prole porque os hemizigotos portadores das letais não sobreviverão (Fig. 13.5).

Mutações com Efeitos Fenotípicos: Geralmente Deletérias e Recessivas

A maioria dos milhares de mutações que foram identificadas e estudadas por geneticistas são deletérias e recessivas. Este resultado deve ser esperado se considerarmos o que é conhecido sobre o controle genético do metabolismo e pelas técnicas disponíveis para identificar mutações. Como discutido no Cap. 11, o metabolismo ocorre por seqüências de reações químicas, cada etapa catalisada por uma enzima específica codificada por um ou mais genes. Mutações nestes genes frequentemente produzem bloqueios em vias

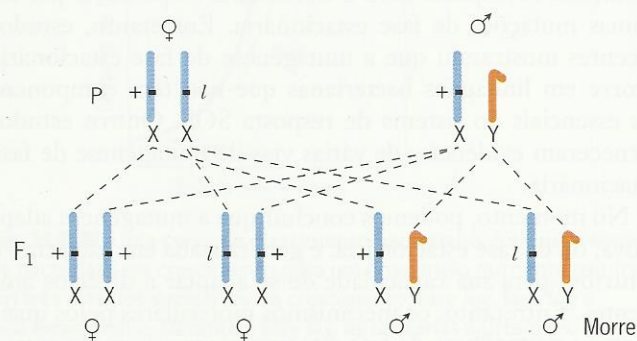


Fig. 13.5 ■ Alteração da proporção de sexos por uma mutação letal recessiva ligada ao X. As fêmeas heterozigotas para um letal recessivo ligado ao X produzem uma prole de fêmeas e machos em uma proporção de 2:1.

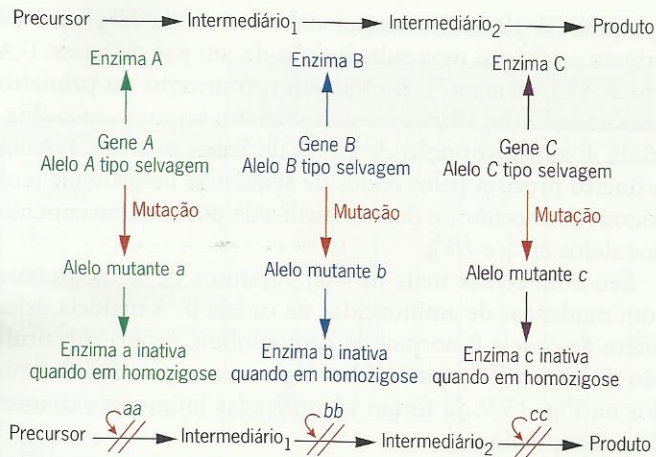


Fig. 13.6 ■ Alelos mutantes recessivos em geral resultam em bloqueios nas vias metabólicas. As vias podem ter apenas algumas etapas, como diagramado aqui, ou ter muitas. O alelo tipo selvagem de cada gene codifica uma enzima funcional que catalisa a reação apropriada. A maioria das mutações que ocorrem nos genes tipo selvagem resultam em formas alteradas da enzima com atividade reduzida ou ausente. No estado homozigoto, os alelos mutantes que produzem produtos inativos causam bloqueios metabólicos (⇒) devido à falta da atividade enzimática requerida.

metabólicas (Fig. 13.6). Tais bloqueios ocorrem porque alterações nas seqüências de pares de bases dos genes em geral causam mudanças nas seqüências de aminoácidos de polipeptídeos (Fig. 13.7), que podem resultar em produtos não funcionais (Fig. 13.6). De fato, este é o efeito mais comumente observado de mutações facilmente detectadas.

Considerando um alelo tipo selvagem que codifica uma enzima ativa e alelos mutantes que codificam enzimas menos ativas ou totalmente inativas, é aparente por que a maioria das mutações observadas seriam recessivas. Se uma célula contém tanto formas ativas quanto inativas de determinada enzima, a forma ativa em geral irá catalisar a reação em questão. Portanto, o alelo que especifica o produto ativo geralmente será dominante e o alelo que codifica o produto inativo será recessivo (Cap. 4).

Devido à redundância e à ordem no código genético (Cap. 12), muitas mutações não têm efeito sobre o fenótipo do organismo; elas são chamadas **mutações neutras**. Mas por que a maioria das mutações com efeitos fenotipicamente reconhecíveis resultam em atividade diminuída do produto gênico ou em nenhuma atividade do produto gênico? Um alelo tipo selvagem de um gene que codifica uma enzima do tipo selvagem ou proteína estrutural terá sido selecionado por atividade ótima no curso da evolução. Assim, as mutações, que causam mudanças aleatórias nas seqüências de aminoácidos altamente adaptadas, em geral darão produtos menos ativos ou totalmente inativos. Você pode fazer uma analogia com qualquer máquina complexa cuidadosamente construída tal como um computador ou um automóvel. Se você modifica aleatoriamente um componente essencial, a máquina improvavelmente funcionará tão bem quanto fazia antes da mudança. Essa visão de mutação e da interação entre alelos mutante e tipo selvagem ajusta-se à observação de que a maioria das mutações com efeitos fenotípicos reconhecíveis são recessivas e deletérias.

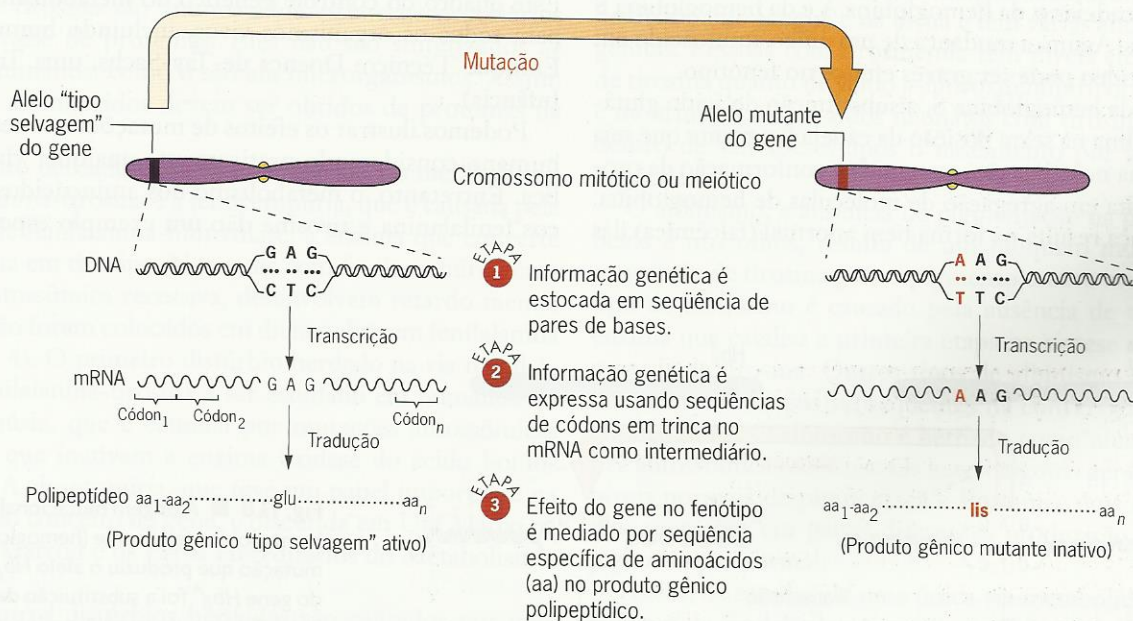


Fig. 13.7 ■ Visão geral do processo de mutação e a expressão dos alelos tipo selvagem e mutante. As mutações alteram as seqüências de pares de nucleotídeos nos genes, os quais, por sua vez, causam mudanças nas seqüências de aminoácidos dos polipeptídeos codificados por estes genes. Um par de bases G-C (acima, à esquerda) mudou para um par de bases A-T (acima, à direita). Esta mutação muda um códon do mRNA de GAG para AAG e um aminoácido no produto polipeptídico de ácido glutâmico (glu) para lisina (lis). Tais mudanças em geral dão produtos gênicos não funcionais.

Efeitos de Mutações em Genes de Globina Humana

As hemoglobinas humanas mutantes fornecem boas ilustrações dos efeitos deletérios da mutação. Nos Caps. 1 e 12, discutimos a estrutura da hemoglobina e os efeitos traumáticos de uma variante de hemoglobina, a hemoglobina falciforme. Lembre que a forma principal da hemoglobina em adultos (hemoglobina A) contém duas **cadeias alfa** (α) idênticas e duas **cadeias beta** (β) idênticas. Cada polipeptídeo α consiste em uma sequência específica de 141 aminoácidos, enquanto cada cadeia β possui 146 aminoácidos. Devido às similaridades em suas sequências de aminoácidos, acredita-se que todas as cadeias de globina (e, portanto, seus genes estruturais) evoluíram de um progenitor comum.

Muitas variantes diferentes da hemoglobina adulta foram identificadas em populações humanas, e várias delas têm graves efeitos fenotípicos. Muitas das variantes foram inicialmente detectadas por seu comportamento eletroforético alterado (movimento em um campo elétrico devido a diferenças de carga – veja Cap. 9). As variantes de hemoglobina fornecem uma excelente ilustração dos efeitos de mutação nas estruturas e funções de produtos gênicos e, finalmente, nos fenótipos das pessoas afetadas.

Quando as sequências de aminoácidos das cadeias β da hemoglobina A e da hemoglobina de pacientes com anemia falciforme (hemoglobina S) foram determinadas e comparadas, foi descoberto que a hemoglobina S difere da hemoglobina A em apenas uma posição. O sexto aminoácido a partir da ponta amino da cadeia β da hemoglobina A é o ácido glutâmico (um aminoácido de carga negativa). A cadeia β da hemoglobina S contém valina (sem carga em pH neutro) nesta posição. As cadeias α da hemoglobina A e da hemoglobina S são idênticas. Assim, a mudança de um único aminoácido em um polipeptídeo pode ter graves efeitos no fenótipo.

No caso da hemoglobina S, a substituição de ácido glutâmico por valina na sexta posição da cadeia β permite que seja formada uma nova ligação, que muda a conformação da proteína e resulta em agregação de moléculas de hemoglobina. Esta mudança resulta na forma bem anormal (falcêmica) das

hemácias. A alteração mutacional no alelo da $Hb\beta^A$ que dá origem a $Hb\beta^S$ foi uma substituição de um par de bases T-A por A-T, com uma T no filamento transcrito no primeiro caso e uma A no filamento transcrito no segundo caso (Fig. 13.8). Esta substituição de pares de bases A-T \rightarrow T-A foi primeiro prevista pelos dados de sequência de proteína e os códons conhecidos, e depois verificada por sequenciamento dos alelos $Hb\beta^A$ e $Hb\beta^S$.

São conhecidas mais de 100 variantes de hemoglobina com mudanças de aminoácidos na cadeia β . A maioria delas difere da cadeia β normal de hemoglobina A pela substituição de um único aminoácido. Alguns exemplos são ilustrados na Fig. 13.9. Já foram identificadas inúmeras variantes do polipeptídeo α .

Os exemplos de hemoglobina mostram que a mutação é um processo no qual mudanças na estrutura do gene, em geral mudanças em um ou alguns pares de bases, podem causar mudanças nas sequências de aminoácidos dos produtos polipeptídicos dos genes. Estas alterações na estrutura das proteínas, por sua vez, causam mudanças no fenótipo que são reconhecidos como mutantes.

Mutações em Humanos: Bloqueios em Vias Metabólicas

No Cap. 11, discutimos o controle genético de vias metabólicas, em que cada etapa em uma via é catalisada por uma enzima codificada por um ou mais genes. Quando ocorrem mutações em tais genes, eles em geral causam bloqueios metabólicos (veja Fig. 13.6) que levam a fenótipos anormais. Este quadro do controle genético do metabolismo é válido para todos os organismos vivos, incluindo humanos (veja Enfoque Técnico: Doença de Tay-Sachs, uma Tragédia na Infância).

Podemos ilustrar os efeitos de mutações no metabolismo humano considerando praticamente qualquer via metabólica. Entretanto, o metabolismo dos aminoácidos aromáticos fenilalanina e tirosina dão um exemplo especialmente

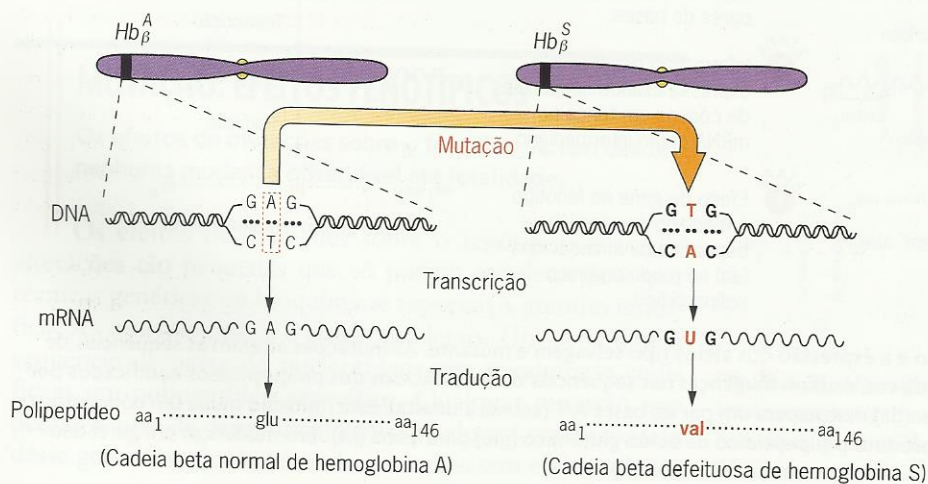


Fig. 13.8 ■ A origem mutacional da hemoglobina falciforme (hemoglobina S). A mutação que produziu o alelo $Hb\beta^S$ a partir do gene $Hb\beta^A$ foi a substituição de um par de bases T-A por um par A-T, onde o filamento inferior do DNA é o filamento transcrito. O resultado é a substituição do aminoácido ácido glutâmico na cadeia β da hemoglobina A pelo aminoácido valina presente na mesma posição na hemoglobina S.

Sequência da cadeia beta de hemoglobina

	Posição do aminoácido														
	1	2	6	7	16	24	26	56	63	95	108	124	141	146	
Hemoglobina A:	Val	His ⁺	Glu ⁻	Glu ⁻	Gli	Gli	Glu ⁻	Gli	His ⁺	Lis ⁺	Asn	Pro	Leu	His ⁺	
Variante de Hb															
Tokuchi		Tir													
Hb S			Val												
Hb C			Lis ⁺												
Hb G				Gli											
J Baltimore					Asp ⁻										
Savannah						Val									
Hb E ⁻							Lis ⁺								
J Bangkok								Asp ⁻							
Zurich									Arg ⁺						
M Saskatoon									Tir						
N Baltimore										Glu ⁻					
Yoshizuka											Asp ⁻				
Khartoum												Arg ⁺			
Olmsted													Arg ⁺		

Fig. 13.9 ■ Algumas substituições de aminoácidos que foram documentadas nas cadeias de β -globina humana. Todos os exemplos mostrados resultaram de substituições de um par de bases no gene de $Hb\beta^A$ (análogo à mostrada para a hemoglobina S na Fig. 13.8). Note que muitas das substituições de aminoácidos envolvem mudanças de carga, muito embora apenas 5 dos 20 aminoácidos comuns tenham uma carga final em pH neutro. Substituições envolvendo mudanças de carga alteram a conformação da proteína mais frequentemente que outras substituições de aminoácidos. Assim, tais substituições mais provavelmente causam mudanças no funcionamento da proteína.

bom porque alguns dos primeiros estudos de mutações em humanos revelaram bloqueios nessa via (Fig. 13.10). Fenilalanina e tirosina são aminoácidos essenciais necessários para a síntese de proteínas. Eles não são sintetizados *de novo* em humanos, como o são em microrganismos. Assim, ambos os aminoácidos devem ser obtidos de proteínas da dieta.

O defeito herdado mais bem conhecido no metabolismo de fenilalanina-tirosina é a fenilcetonúria, que é causada pela ausência de fenilalanina-hidroxilase, a enzima que converte fenilalanina em tirosina. Neonatos com fenilcetonúria, uma doença autossômica recessiva, desenvolvem retardo mental grave se não forem colocados em dieta pobre em fenilalanina (veja Cap. 4). O primeiro distúrbio herdado na via metabólica de fenilalanina-tirosina a ser estudado em humanos foi a alcaptonúria, que é causada por mutações autossômicas recessivas que inativam a enzima oxidase do ácido homogentísico. A alcaptonúria, que teve um papel importante na evolução do conceito de gene, é discutida em Um Marco na Genética: Garrod e os Erros Hereditários do Metabolismo, no Cap. 4.

Dois outros distúrbios herdados são causados por mutações em genes que codificam enzimas necessárias para o catabolismo da tirosina, ambos herdados como autossômicos recessivos. Tirosinose e tirosinemia resultam da falta das enzimas tirosina-transaminase e oxidase do ácido *p*-hidroxifenilpirúvico, respectivamente (Fig. 13.10). A tirosinose é

muito rara; apenas alguns casos foram estudados. Pessoas com tirosinose apresentam aumentos acentuados nos níveis de tirosina no sangue e na urina e têm várias anomalias congênitas. Pessoas com tirosinemia têm níveis elevados tanto de tirosina quanto de ácido *p*-hidroxifenilpirúvico no sangue e na urina. A maioria dos neonatos com tirosinemia morre dentro de seis meses após o nascimento por insuficiência hepática.

O albinismo, a ausência de pigmentação na pele, nos cabelos e nos olhos, resulta de um bloqueio mutacional na conversão de tirosina para o pigmento escuro melanina. Um tipo de albinismo é causado pela ausência de tirosinase, a enzima que catalisa a primeira etapa na síntese de melanina a partir de tirosina. Outros tipos de albinismo resultam de bloqueios em etapas subsequentes na conversão de tirosina em melanina. O albinismo é herdado como uma característica autossômica recessiva. Os heterozigotos geralmente têm níveis normais de pigmentação. Portanto, dois albinos que têm mutações em genes diferentes produzirão filhos com pigmentação normal.

Assim, os estudos de uma única via metabólica, do metabolismo de fenilalanina-tirosina, revelaram cinco distúrbios herdados diferentes, todos causados por mutações em genes que controlam etapas nesta via. Exemplos similares do controle genético do metabolismo podem ser obtidos examinando-se essencialmente qualquer outra via metabólica em humanos.



ENFOQUE TÉCNICO: Doença de Tay-Sachs, uma Tragédia na Infância

De todos os distúrbios humanos herdados, um dos mais trágicos é a doença de Tay-Sachs. As crianças homozigotas para o gene mutante que causa a doença de Tay-Sachs são normais ao nascimento. Entretanto, dentro de alguns meses, tornam-se hipersensíveis a barulhos altos e desenvolvem um ponto vermelho-cereja na retina do olho. Estes primeiros sintomas da doença em geral não são detectados pelos pais e médicos. De seis meses a um ano após o nascimento, crianças com Tay-Sachs começam a sofrer uma degeneração neurológica progressiva que rapidamente resulta em retardo mental, cegueira, surdez e perda geral do controle das funções corpóreas. Aos dois anos de idade, elas em geral estão totalmente paralisadas e desenvolvem infecções respiratórias crônicas. A morte geralmente ocorre aos três a quatro anos de idade.

Embora o defeito molecular responsável pela doença de Tay-Sachs seja conhecido, não há tratamento eficaz para o distúrbio. O único aspecto positivo da doença de Tay-Sachs é que é rara na maioria das populações. Entretanto, isto é de pouco conforto para os judeus ashkenazi da Europa Central e seus descendentes. A doença de Tay-Sachs ocorre em cerca de 1 em 3.600 de seus filhos, e cerca de 1 em cada 30 adultos nestas populações judaicas portam o gene mutante no estado heterozigoto. Se duas pessoas destas populações se casam, a chance de que ambas portem o gene mutante é de cerca de 1 em 1.000 ($0,033 \times 0,033$). Se ambos forem portadores, em média, um quarto de seus filhos serão homozigotos para o gene mutante e desenvolverão a doença de Tay-Sachs.

A mutação que causa a doença de Tay-Sachs está situada no gene *HEXA*, que codifica a enzima hexosaminidase A. Esta enzima atua em um lipídeo complexo chamado de gangliosídeo G_{M2} , clivando-o em um gangliosídeo menor (G_{M3}) e *N*-acetil-*D*-galactosamina, como mostrado na Fig. 1.

A função do gangliosídeo G_{M2} é revestir as células nervosas, isolando-as de eventos que ocorrem nas células vizinhas, acelerando assim a transmissão de impulsos nervosos. Na ausência da enzima que o quebra, o gangliosídeo G_{M2} acumula-se e literalmente encobre as células nervosas. Esse acúmulo de lipídeos complexos nos neurônios bloqueia sua ação, provocando degeneração do sistema nervoso e finalmente paralisia.

Embora a doença de Tay-Sachs tenha sido descrita por Warren Tay em 1881 e a base bioquímica seja conhecida há mais de 20 anos, ainda não há um tratamento eficaz para este trágico distúrbio. Embora alguns distúrbios herdados possam ser tratados por terapia enzimática, administrando a enzima ausente aos pacien-

tes, este enfoque não é factível na doença de Tay-Sachs porque a enzima não penetrará na barreira que separa as células cerebrais do sistema circulatório. Além disso, a terapia gênica de células somáticas, fornecendo cópias funcionais do gene afetado a células somáticas (Cap. 17), no momento não é possível porque não há procedimento estabelecido para introduzir genes em neurônios. De fato, os cientistas ainda não sabem quais células nervosas são responsáveis pela degeneração neurológica que ocorre em crianças com a doença.

A doença de Tay-Sachs pode ser detectada pré-natalmente por amniocentese (Cap. 6), e este procedimento tem sido amplamente usado para diagnosticar o distúrbio. Recentemente, foi desenvolvido um teste sensível de DNA que permite aos cientistas detectar o gene mutante que causa a doença de Tay-Sachs no DNA isolado de uma única célula (Cap. 17). Este teste de DNA foi usado para triar pré-embriões de oito células produzidos por fertilização *in vitro* para a mutação Tay-Sachs. Uma célula é usada para o teste de DNA e as outras sete conservam a capacidade de se desenvolver em um embrião normal quando implantadas no útero da mãe. Apenas embriões que dão um resultado normal – os que não são homozigotos para o gene mortal Tay-Sachs – são implantados. Este procedimento permite a genitores que são ambos portadores do gene mutante ter filhos sem se preocupar com o nascimento de uma criança com doença de Tay-Sachs.

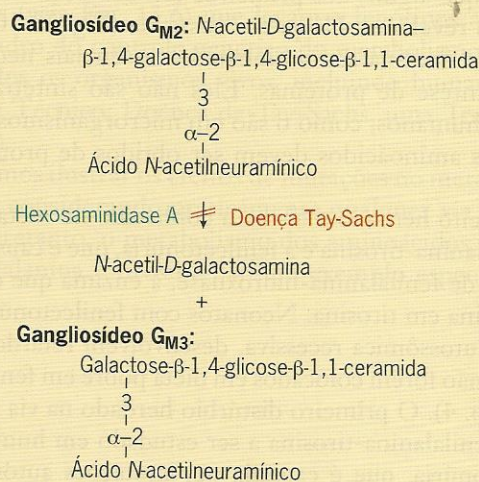


Fig. 1 ■ O defeito metabólico em humanos na doença de Tay-Sachs.

Mutações Letais Condicionais: Instrumentos Poderosos para Estudos Genéticos

De todas as mutações, de isoalelos a letais, as **mutações letais condicionais** são as mais úteis para estudos genéticos. Constituem mutações que são (1) letais em um ambiente,

a **condição restritiva**, mas são (2) viáveis em um segundo ambiente, a **condição permissiva**. As mutações letais condicionais permitem aos geneticistas identificar e estudar mutações em genes essenciais que resultam em perda completa de atividade do produto gênico, mesmo em organismos haplóides. Os mutantes que levam letais condicionais podem ser propagados em condições permissivas, e a informação sobre

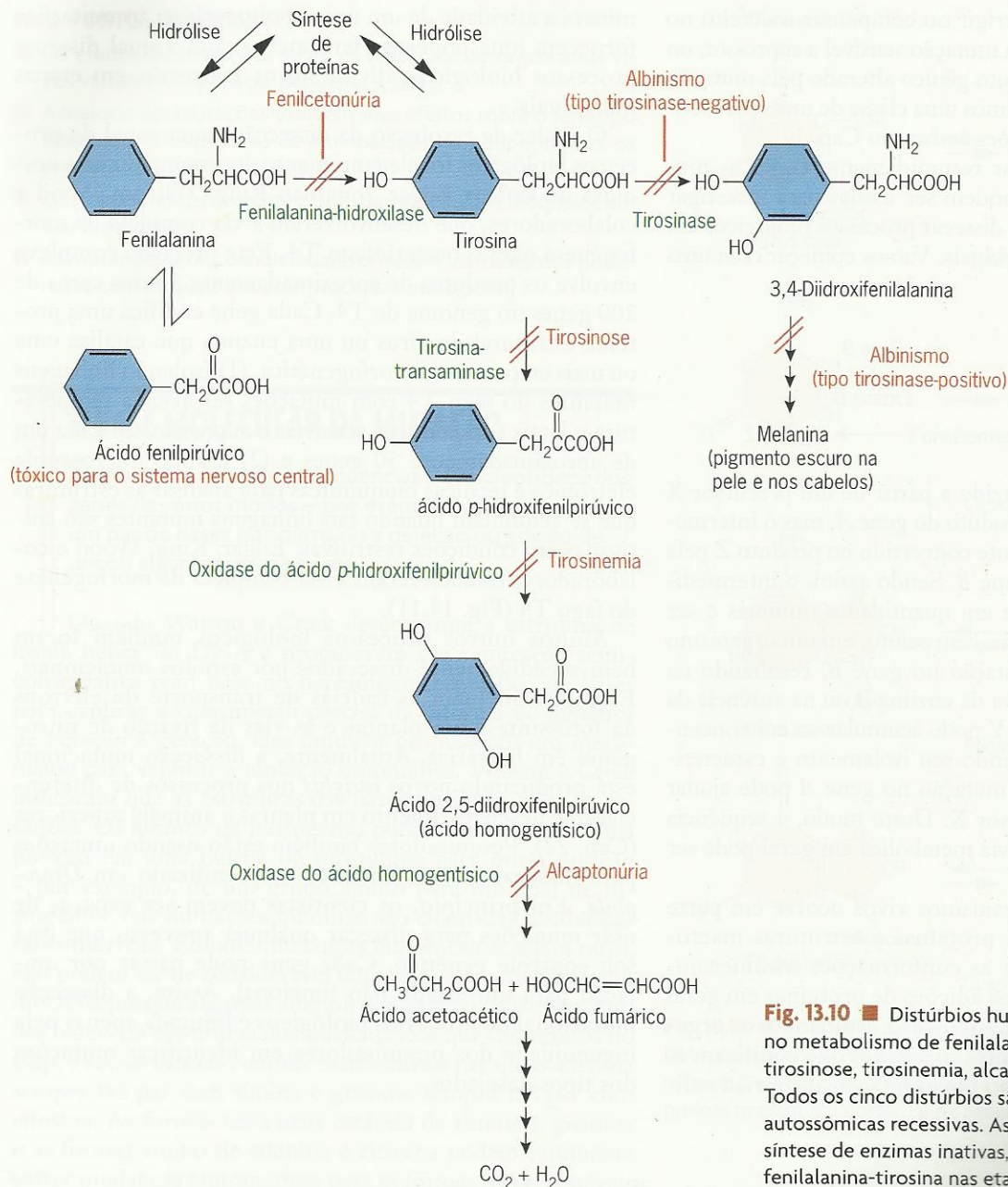


Fig. 13.10 ■ Distúrbios humanos herdados com defeitos no metabolismo de fenilalanina-tirosina: fenilcetonúria, tirosinose, tirosinemia, alcaptonúria e albinismo. Todos os cinco distúrbios são causados por mutações autossômicas recessivas. As mutações, que resultam na síntese de enzimas inativas, bloqueiam o metabolismo de fenilalanina-tirosina nas etapas indicadas.

as funções dos produtos gênicos podem ser deduzidas estudando-se as consequências de sua ausência nas condições restritivas. Mutações letais condicionais foram usadas para investigar uma grande gama de processos biológicos, desde o desenvolvimento até a fotossíntese.

As três classes principais de mutantes com fenótipos letais condicionais são (1) mutantes auxotróficos, (2) mutantes sensíveis a temperatura e (3) mutantes sensíveis a supressores. **Auxotróficos** são mutantes incapazes de sintetizar um metabólito essencial (aminoácido, purina, pirimidina, vitamina, e assim em diante) que é sintetizado por organismos tipo selvagem ou *prototróficos* da mesma espécie. Os auxotróficos crescerão e se reproduzirão quando o metabólito é fornecido no meio (a condição permissiva). Eles não crescerão quando

o metabólito essencial está ausente (a condição restritiva). **Mutantes sensíveis a temperatura** crescerão em uma temperatura e não em outra. A maioria dos mutantes sensíveis a temperatura são sensíveis ao calor. Entretanto, alguns são sensíveis ao frio. A sensibilidade a temperatura geralmente resulta de aumento de labilidade a calor ou frio do produto gênico mutante – por exemplo, uma enzima que é ativa em temperatura baixa mas parcial ou totalmente inativa em temperaturas mais altas. Ocasionalmente, apenas a síntese do produto gênico é sensível a temperatura, e, uma vez feito, o produto do gene mutante pode ser tão estável quanto o produto gênico tipo selvagem. Os **mutantes sensíveis a supressor** são viáveis quando um segundo fator genético, um supressor, está presente, mas são inviáveis na ausência de supressor.

O gene supressor pode corrigir ou compensar o defeito no fenótipo que é causado pela mutação sensível a supressor, ou pode fazer com que o produto gênico alterado pela mutação não seja essencial. Já discutimos uma classe de mutações sensíveis a supressor, as mutações *âmbar*, no Cap. 12.

Agora, vamos considerar resumidamente como as mutações letais condicionais podem ser usadas para investigar processos biológicos – para dissecar processos biológicos em suas partes, ou etapas, individuais. Vamos começar com uma via biossintética simples:



O intermediário Y é produzido a partir de um precursor X pela ação da enzima A, o produto do gene A, mas o intermediário Y pode ser rapidamente convertido no produto Z pela enzima B, o produto do gene B. Sendo assim, o intermediário Y pode estar presente em quantidades mínimas e ser difícil de isolar e caracterizar. Entretanto, em um organismo mutante que tem uma mutação no gene B, resultando na síntese de uma forma inativa da enzima B ou na ausência da enzima B, o intermediário Y pode acumular-se em concentrações muito altas, facilitando seu isolamento e caracterização. Similarmente, uma mutação no gene A pode ajudar na identificação do precursor X. Deste modo, a sequência de etapas em determinada via metabólica em geral pode ser determinada.

A morfogênese em organismos vivos ocorre em parte pela adição sequencial de proteínas a estruturas macromoleculares para produzir as conformações tridimensionais finais, e a sequência das adições de proteínas em geral pode ser determinada isolando-se e estudando-se organismos mutantes com defeitos nos genes que codificam as proteínas envolvidas. Como uma mutação apropriada eli-

minará a atividade de um único polipeptídeo, as mutações fornecem uma poderosa ferramenta com a qual dissecar processos biológicos, dividindo os processos em etapas individuais.

O poder de resolução da dissecação mutacional de processos biológicos foi elegantemente documentado pela pesquisa de Robert Edgar, Jonathan King, William Wood e colaboradores, que desenvolveram a via completa da morfogênese para o bacteriófago T4. Este processo complexo envolve os produtos de aproximadamente 50 dos cerca de 200 genes no genoma de T4. Cada gene codifica uma proteína estrutural do vírus ou uma enzima que catalisa uma ou mais etapas na via morfogenética. (1) Isolando linhagens mutantes do fago T4 com mutações sensíveis a temperatura e letais condicionais sensíveis a supressor em cada um de aproximadamente 50 genes e (2) usando microscopia eletrônica e técnicas bioquímicas para analisar as estruturas que se acumulam quando tais linhagens mutantes são cultivadas em condições restritivas, Edgar, King, Wood e colaboradores estabeleceram a via completa da morfogênese do fago T4 (Fig. 13.11).

Muitos outros processos biológicos também foram bem-sucedidamente dissecados por estudos mutacionais. Exemplos incluem as cadeias de transporte de elétrons da fotossíntese em plantas e as vias de fixação de nitrogênio em bactérias. Atualmente, a dissecação mutacional está produzindo novos *insights* nos processos de diferenciação e desenvolvimento em plantas e animais superiores (Cap. 22). Pesquisadores também estão usando mutações para dissecar comportamento e aprendizado em *Drosophila*. Em princípio, os cientistas devem ser capazes de usar mutações para dissecar qualquer processo que está sob controle genético. Cada gene pode passar por mutação para um estado não funcional. Assim, a dissecação mutacional de processos biológicos é limitada apenas pela ingenuidade dos pesquisadores em identificar mutações dos tipos desejados.

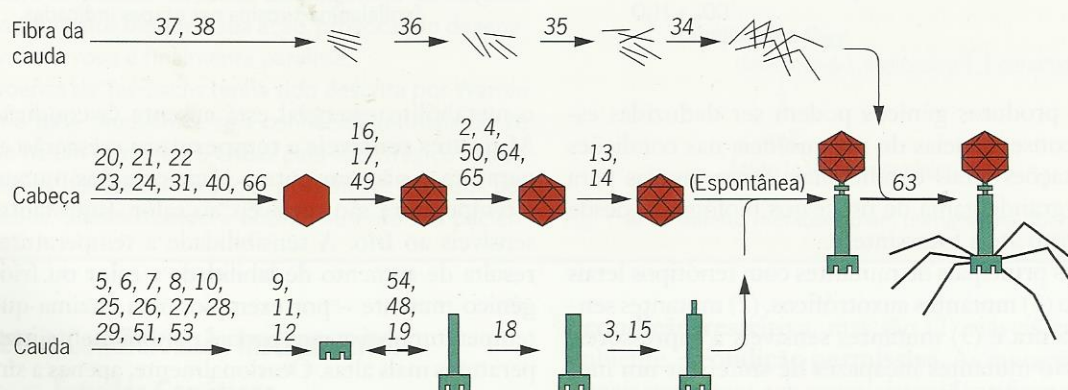


Fig. 13.11 ■ Mapa abreviado de morfogênese no bacteriófago T4. A cabeça, a cauda e as fibras da cauda são produzidas por ramos separados da via e são então unidas nos estágios finais da morfogênese. Os números identificam os genes de T4 cujos produtos são necessários em cada etapa na via. As sequências das primeiras etapas na formação da cabeça e da cauda são conhecidas, mas foram omitidas aqui para manter o diagrama conciso.

PONTOS IMPORTANTES

- Os efeitos de mutações sobre os fenótipos de organismos vivos variam de mudanças pequenas a letais.
- A maioria das mutações exercem seus efeitos sobre o fenótipo alterando as seqüências de aminoácidos de polipeptídeos, os produtos primários dos genes.
- Os polipeptídeos mutantes, por sua vez, causam bloqueios em vias metabólicas.
- As mutações letais condicionais fornecem instrumentos poderosos com os quais dissecar processos biológicos.

A BASE MOLECULAR DA MUTAÇÃO

As mutações alteram as seqüências de nucleotídeos dos genes de vários modos – por exemplo, a substituição de um par de bases por outro ou a deleção ou adição de um ou alguns pares de bases.

Quando Watson e Crick descreveram a estrutura de dupla hélice do DNA e propuseram sua replicação semi-conservativa com base no pareamento específico de bases para explicar a transmissão precisa da informação genética de geração a geração, eles também propuseram um mecanismo para explicar a mutação espontânea. Watson e Crick indicaram que as estruturas das bases no DNA não são estáticas. Os átomos de hidrogênio podem mover-se de uma posição em uma purina ou pirimidina para outra posição – por exemplo, de um grupo amino para um anel de nitrogênio. Tais flutuações químicas são chamadas **mudanças tautoméricas**. Embora mudanças tautoméricas sejam raras, elas podem ser de considerável importância no metabolismo do DNA porque algumas alteram o potencial de pareamento das bases. As estruturas dos nucleotídeos que discutimos no Cap. 9 são as formas comuns mais estáveis nas quais adenina sempre faz par com timina e guanina sempre faz par com citosina. As formas ceto mais estáveis de timina e guanina e as formas amino de adenina e citosina podem raramente sofrer mudanças tautoméricas para as formas menos estáveis enol e imino, respectivamente (Fig. 13.12). Seria esperado que as bases existissem em suas formas tautoméricas menos estáveis por apenas curtos períodos de tempo. Entretanto, se uma base existe na forma rara no momento em que está sendo replicada ou sendo incorporada a uma cadeia de DNA nascente, resultará uma mutação. Quando as bases estão presentes em seus raros estados imino ou enol, elas podem formar pares de bases adenina-citosina e guanina-timina (Fig. 13.13a). O efeito final de tal evento, e a subsequente replicação requerida para segregar o par de bases pareado errado, é uma substituição A-T para G-C ou G-C para A-T (Fig. 13.13b).

Mutações que resultam de mudanças tautoméricas nas bases do DNA envolvem a substituição de uma purina em um filamento do DNA por outra purina e a substituição de uma pirimidina no filamento complementar por outra piri-

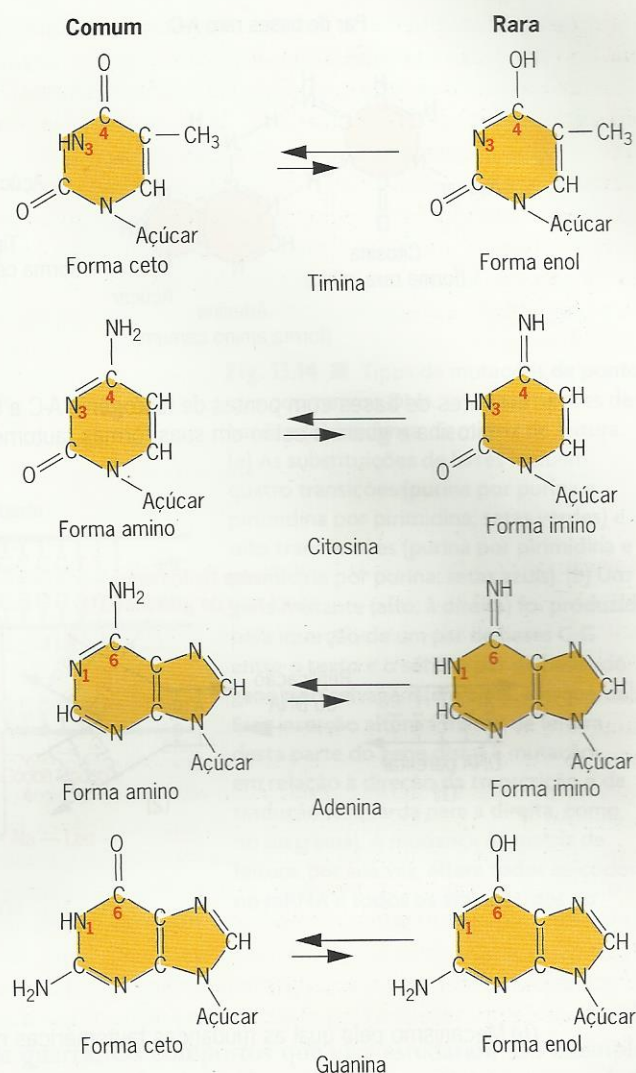
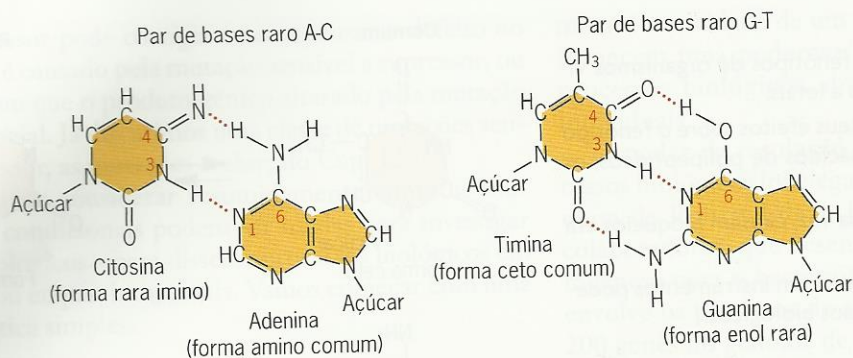
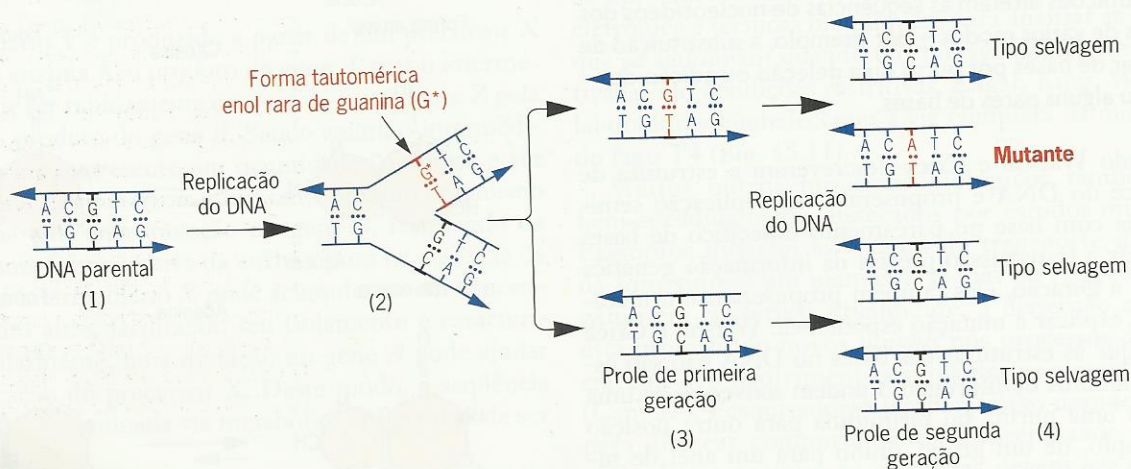


Fig. 13.12 ■ Formas tautoméricas das quatro bases comuns no DNA. As mudanças nos átomos de hidrogênio entre as posições de número 3 e 4 das pirimidinas e entre as posições de número 1 e 6 das purinas mudam seu potencial de pareamento de bases.

midina. Tais substituições de pares de bases são chamadas **transições**. Substituições de pares de bases envolvidas na substituição de uma purina por uma pirimidina e vice-versa são chamadas **transversões**. Existem três substituições, uma transição e duas transversões, possíveis para cada par de bases. São possíveis um total de quatro transições diferentes e oito transversões diferentes (Fig. 13.14a). Outro tipo de mutação de ponto envolve a adição ou deleção de um ou alguns pares de bases. Adições e deleções de pares de bases são coletivamente chamadas **mudanças de matriz de leitura** porque alteram a matriz de leitura de todas as trincas de pares de bases (trincas de DNA que especificam códons no mRNA e aminoácidos no produto polipeptídico do gene) nos genes que estão distais ao sítio no qual ocorreu a mutação (Fig. 13.14b).



(a) Pares de bases com pontes de hidrogênio A-C e G-T que se formam quando citosina e guanina estão em suas formas tautoméricas raras imino e enol.



(b) Mecanismo pelo qual as mudanças tautoméricas nas bases do DNA causam mutações.

Fig. 13.13 ■ Os efeitos das mudanças tautoméricas nos nucleotídeos no DNA em (a) pareamento de bases e (b) mutação. Raros pares de bases A-C e G-T como os mostrados em (a) também se formam quando timina e adenina estão em suas raras formas enol e imino, respectivamente. (b) Uma guanina (1) sofre uma mudança tautomérica para sua rara forma enol (G^*) na época da replicação (2). Na sua forma enol, guanina faz par com timina (2). Durante a replicação subsequente (3 para 4), a guanina muda de volta à sua forma ceto mais estável. A timina incorporada oposta à forma enol da guanina (2) dirige a incorporação de adenina durante a replicação seguinte (3 para 4). O resultado final é uma substituição de par de bases G-C para A-T.

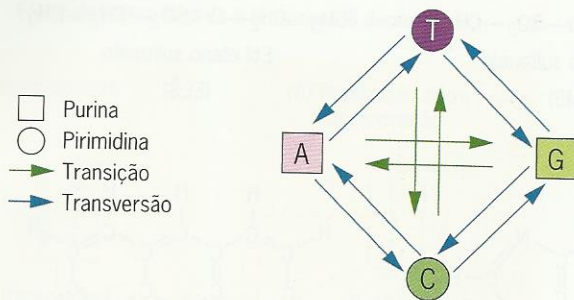
Todos os três tipos de mutações de ponto – transições, transversões e mudanças de matriz de leitura – estão presentes entre mutações de ocorrência espontânea. Uma surpreendente grande proporção de mutações espontâneas que foram estudadas são adições e deleções de um único par de bases, e não substituições de pares de bases. Estas mudanças de matriz de leitura quase sempre resultam na síntese de produtos gênicos protéicos não funcionais.

Embora ainda reste muito a ser aprendido sobre as causas, os mecanismos moleculares e a frequência das mutações espontâneas, os três fatores principais são (1) a precisão da maquinaria de replicação do DNA, (2) a eficiência dos mecanismos que se desenvolveram para reparo do DNA danificado e (3) o grau de exposição a agentes mutagênicos presentes no ambiente. Perturbações do aparelho de replicação do DNA

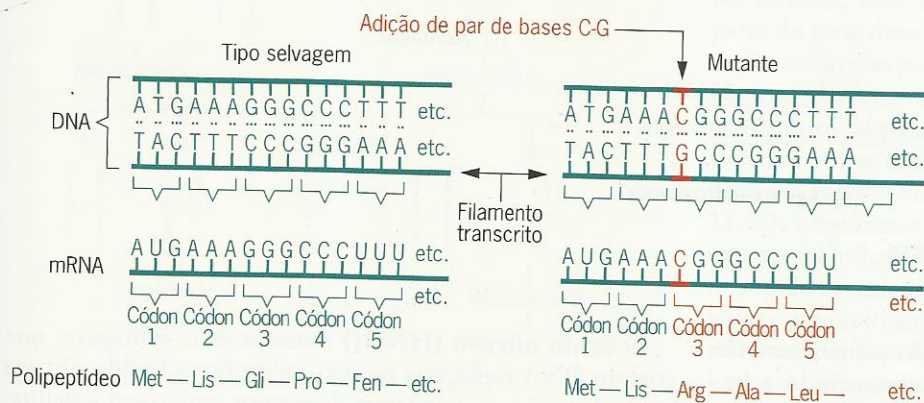
ou de sistemas de reparo do DNA, ambos sob controle genético, mostraram-se como causadoras de grandes aumentos nas taxas de mutação.

Mutações Induzidas

Muitas mutações de ocorrência natural foram identificadas e estudadas pelos primeiros geneticistas. Entretanto, a ciência da genética mudou muito em 1927 quando Hermann J. Muller descobriu que os raios X induziam mutações em *Drosophila*. A capacidade de induzir mutações abriu as portas para um enfoque totalmente novo da análise genética. Geneticistas agora podem induzir mutações em genes de interesse e então estudar os efeitos da falta dos produtos gênicos. Dis-



(a) Podem ocorrer doze substituições diferentes de bases no DNA.



(b) Inserções ou deleções de um ou dois pares de bases alteram a matriz de leitura do gene distal ao sítio da mutação.

Fig. 13.14 ■ Tipos de mutações de ponto que ocorrem no DNA: (a) substituições de bases e (b) mudança de matriz de leitura. (a) As substituições de bases incluem quatro transições (purina por purina e pirimidina por pirimidina; setas verdes) e oito transversões (purina por pirimidina e pirimidina por purina; setas azuis). (b) Um gene mutante (alto, à direita) foi produzido pela inserção de um par de bases C-G entre o sexto e o sétimo par de bases do gene tipo selvagem (em cima, à esquerda). Essa inserção altera a matriz de leitura desta parte do gene distal à mutação, em relação à direção da transcrição e da tradução (esquerda para a direita, como no diagrama). A mudança na matriz de leitura, por sua vez, altera todos os códons no mRNA e todos os aminoácidos no

cutimos a engenhosa demonstração de Muller de mutações induzidas por raios X no cromossomo X de *Drosophila* em Um Marco na Genética: Muller Demonstrou que os Raios X São Mutagênicos. Ele recebeu o Prêmio Nobel de Fisiologia ou Medicina em 1946 por esta descoberta. Trabalhos subsequentes mostraram que os raios X são mutagênicos em todos os organismos e que muitos outros agentes – físicos, químicos e elementos genéticos de transposição – são similarmente mutagênicos.

Os raios X têm muitos efeitos sobre tecidos vivos. Portanto, mutações induzidas por raios X dão poucas informações sobre os mecanismos moleculares pelos quais são produzidas mutações. A descoberta de mutágenos químicos com efeitos específicos sobre o DNA levaram a melhor compreensão do processo de mutação em nível molecular.

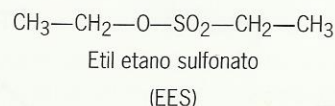
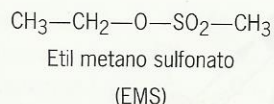
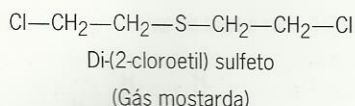
O gás mostarda (mostarda sulfurada) foi a primeira substância química demonstrada como mutagênica. Charlotte Auerbach e seus colaboradores descobriram os efeitos mutagênicos do gás mostarda e compostos correlatos durante a Segunda Guerra Mundial. Entretanto, devido ao uso potencial do gás mostarda na guerra química, o governo britânico colocou seus resultados em uma lista secreta. Assim, Auerbach e colaboradores não puderam nem publicar seus resultados nem discuti-los com outros geneticistas até o fim

da guerra. Os compostos que eles estudaram são exemplos de uma grande classe de mutágenos químicos que transferem grupos alquila (CH_3^- , CH_3CH_2^- , etc.) para as bases no DNA. Assim, eles são chamados de agentes alquilantes. Como os raios X, o gás mostarda exerce muitos efeitos sobre o DNA. Mais tarde, foram descobertos mutágenos químicos que têm efeitos específicos sobre o DNA (Fig. 13.15).

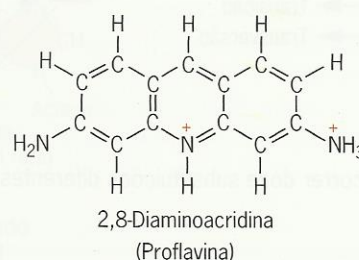
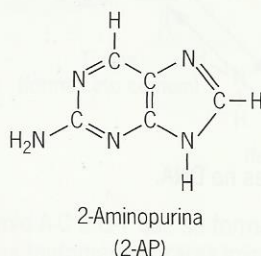
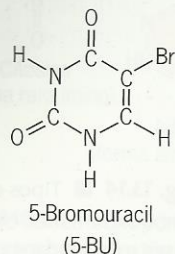
Mutações Induzidas por Substâncias Químicas

Os mutágenos químicos podem ser divididos em dois grupos: (1) os que são mutagênicos para o DNA replicante e não replicante, tais como os agentes alquilantes e o ácido nitroso; e (2) os que são mutagênicos apenas para o DNA replicante, tais como análogos de bases – purinas e pirimidinas com estruturas similares às bases normais no DNA. Os análogos de bases devem ser incorporados às cadeias de DNA no lugar das bases normais durante replicação para exercer seus efeitos mutagênicos. O segundo grupo de mutágenos também inclui os corantes acridina, que se intercalam no DNA e aumentam a probabilidade de erros durante a replicação.

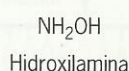
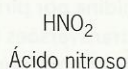
Os **análogos de bases** mutagênicos têm estruturas similares às bases normais e são incorporados ao DNA durante



(a) Agentes alquilantes



(b) Análogos de bases



(d) Agente desaminante

(e) Agente hidroxilante

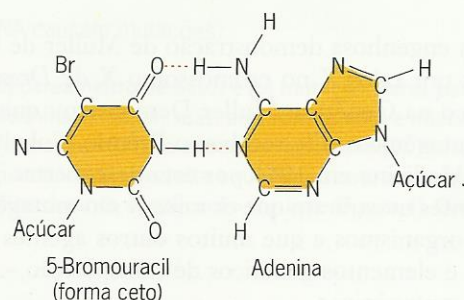
(c) Acridinas

Fig. 13.15 ■ Alguns potentes mutágenos químicos.

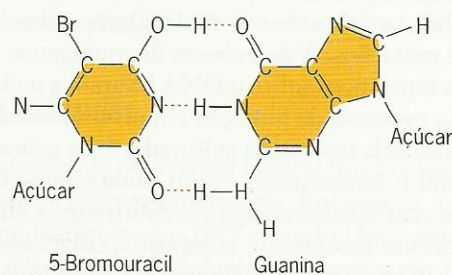
replicação. Entretanto, suas estruturas são suficientemente diferentes das bases normais no DNA aumentando a frequência de pareamentos errados, e assim mutação, durante a replicação. Os dois análogos de bases mais comumente usados são 5-bromouracil e 2-aminopurina. A pirimidina 5-bromouracil é um análogo de timina. O bromo na posição 5 é similar em muitos aspectos ao grupo metila ($-\text{CH}_3$) na posição 5 na timina. Entretanto, o bromo nesta posição muda a distribuição de carga e aumenta a frequência de mudanças tautoméricas (veja Fig. 13.12). Na sua forma ceto mais estável, 5-bromouracil faz par com adenina. Após uma mudança tautomérica para sua forma enol, 5-bromouracil faz par com guanina (Fig. 13.16). O efeito mutagênico da 5-bromouracil é o mesmo que o previsto para mudanças tautoméricas em bases normais (veja Fig. 13.13b), ou seja, transições.

Se 5-bromouracil estiver presente em sua forma enol menos frequente como um trifosfato de nucleosídeo na época de sua incorporação a um filamento nascente de DNA, ela será incorporada em oposição a guanina no filamento molde e irá causar uma transição $\text{G-C} \rightarrow \text{A-T}$ (Fig. 13.17a). Se, entretanto, 5-bromouracil for incorporada em sua forma ceto mais frequente em oposição a adenina (em lugar de timina) e sofrer uma mudança tautomérica para sua forma enol durante uma replicação subsequente, ela causará uma transição $\text{A-T} \rightarrow \text{G-C}$ (Fig. 13.17b). Assim, 5-bromouracil induz transições em ambos os sentidos, $\text{A-T} \leftrightarrow \text{G-C}$. Uma consequência importante da bidirecionalidade de transições induzidas por 5-bromouracil é que mutações originalmente induzidas por este análogo de timina também podem induzir uma mutação reversa para o tipo selvagem com 5-bromouracil. 2-Aminopurina atua de modo similar mas é incorporada em lugar de adenina ou guanina.

O ácido nitroso (HNO_2) é um potente mutágeno que atua no DNA replicante ou não replicante. O ácido nitroso causa desaminação oxidativa de grupos amino em adenina, guanina e citosina. Esta reação converte os grupos amino em grupos ceto e muda o potencial de pontes de hidrogênio das bases modificadas (Fig. 13.18). A adenina é desaminada



(a) 5-Bromouracil: par de bases com adenina.



(b) 5-Bromouracil: par de bases com guanina.

Fig. 13.16 ■ Pareamento de bases entre 5-bromouracil e (a) adenina ou (b) guanina.

Efeito da forma enol de 5-bromouracil durante:

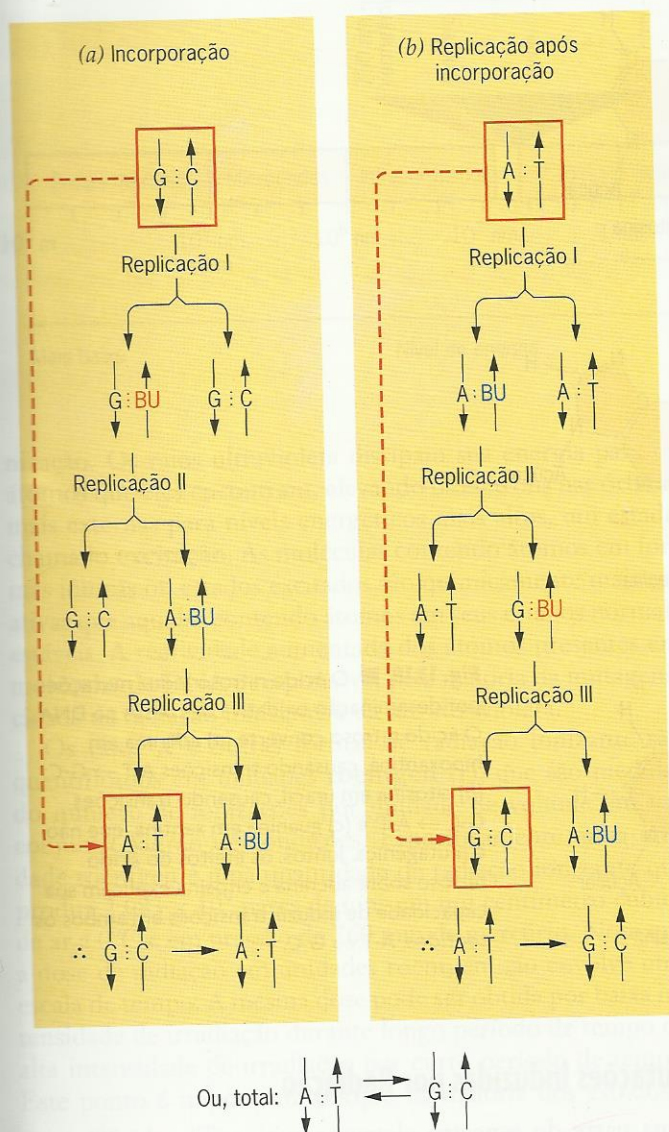


Fig. 13.17 ■ Os efeitos mutagênicos de 5-bromouracil. (a) Quando 5-bromouracil (BU) está presente em sua forma enol menos frequente (laranja) na época da incorporação ao DNA, ela induz transições G-C \rightarrow A-T. (b) Quando 5-bromouracil é incorporada ao DNA em sua forma ceto mais comum (azul) e muda para sua forma enol durante uma replicação subsequente, ela induz transições A-T \rightarrow G-C. Assim, 5-bromouracil pode induzir transições em ambos os sentidos, A-T \leftrightarrow G-C.

em hipoxantina, que faz par de bases com citosina e não com timina. Citosina é convertida em uracil, que faz par de bases com adenina em vez de guanina. A desaminação de guanina produz xantina, mas xantina, como a guanina, faz par de bases com citosina. Assim, a desaminação de guanina não é mutagênica. Como a desaminação de adenina resulta em transições A-T \rightarrow G-C e a desaminação de citosina produz transições G-C \rightarrow A-T, o ácido nitroso induz transições em ambos os sentidos – A-T \leftrightarrow G-C. Como resultado, muta-

ções induzidas por ácido nitroso também induzem retromutação para o tipo selvagem pelo ácido nitroso.

Os **corantes acridina** tais como proflavina (Fig. 13.15), acridina-laranja e toda uma série de compostos correlatos são mutágenos potentes que induzem mudanças de matriz de leitura (veja Fig. 13.14b). As acridinas de carga positiva intercalam-se entre os pares de bases empilhados no DNA (Fig. 13.19). Desta forma, elas aumentam a rigidez e alteram a conformação da dupla hélice, causando inclinações ou dobras na molécula. Quando moléculas de DNA contendo acridinas intercaladas se replicam, ocorrem adições e deleções de um a três pares de bases. Como podemos esperar, essas pequenas adições e deleções, geralmente de um único par de bases, resultam em matrizes de leitura alteradas para a parte do gene distal à mutação (veja Fig. 13.14b). Assim, mutações induzidas por acridina em geral resultam em produtos gênicos não funcionais.

Agentes alquilantes são substâncias que doam grupos alquila a outras moléculas. Eles incluem a mostarda nitrogenada e o metil e etil metano sulfonato (MMS e EMS) (Fig. 13.15), substâncias que têm vários efeitos sobre o DNA. Os agentes alquilantes induzem todos os tipos de mutações, incluindo transições, transversões, mudanças de matriz de leitura e mesmo anomalias cromossômicas, com frequências relativas que dependem da reatividade do agente envolvido. Um mecanismo de mutagenese por agentes alquilantes envolve a transferência de grupos metila ou etila para as bases, resultando em potenciais de pareamento de bases alterados. Por exemplo, EMS causa etilação das bases no DNA nas posições 7-N e 6-O. Quando é produzida 7-etilguanina, ela faz pares de bases com timina para causar transições G-C \rightarrow A-T. Outros produtos de alquilação de bases ativam processos de reparo do DNA propensos a erro que introduzem transições, transversões e mudanças de matriz de leitura durante o processo de reparo. Alguns agentes alquilantes, particularmente agentes alquilantes disfuncionais (aqueles com dois grupos alquila reativos), fazem ligações cruzadas de filamentos ou moléculas de DNA e induzem quebras cromossômicas, que resultam em vários tipos de anomalias cromossômicas (Cap. 6). Os agentes alquilantes como uma classe exibem, portanto, menos efeitos mutagênicos específicos do que os análogos de bases, ácido nitroso ou acridinas.

Em contraste com a maioria dos agentes alquilantes, o **agente hidroxilante** hidroxilamina (NH_2OH) exerce um efeito mutagênico específico. Ele induz apenas transições G-C \rightarrow A-T. Quando o DNA é tratado com hidroxilamina, o grupo amino da citosina é hidroxilado. A hidroxilaminacitosina resultante faz par de bases com adenina, resultando em transições G-C \rightarrow A-T. Devido a sua especificidade, a hidroxilamina tem sido muito útil em classificar mutações de transição. As mutações que são induzidas para reverter ao tipo selvagem pelo ácido nitroso ou análogos de bases, e portanto foram originalmente causadas por transições, podem ser divididas em duas classes com base em sua reversibilidade com hidroxilamina. (1) Aquelas com um par de bases A-T no sítio mutante não serão induzidas para reverter por

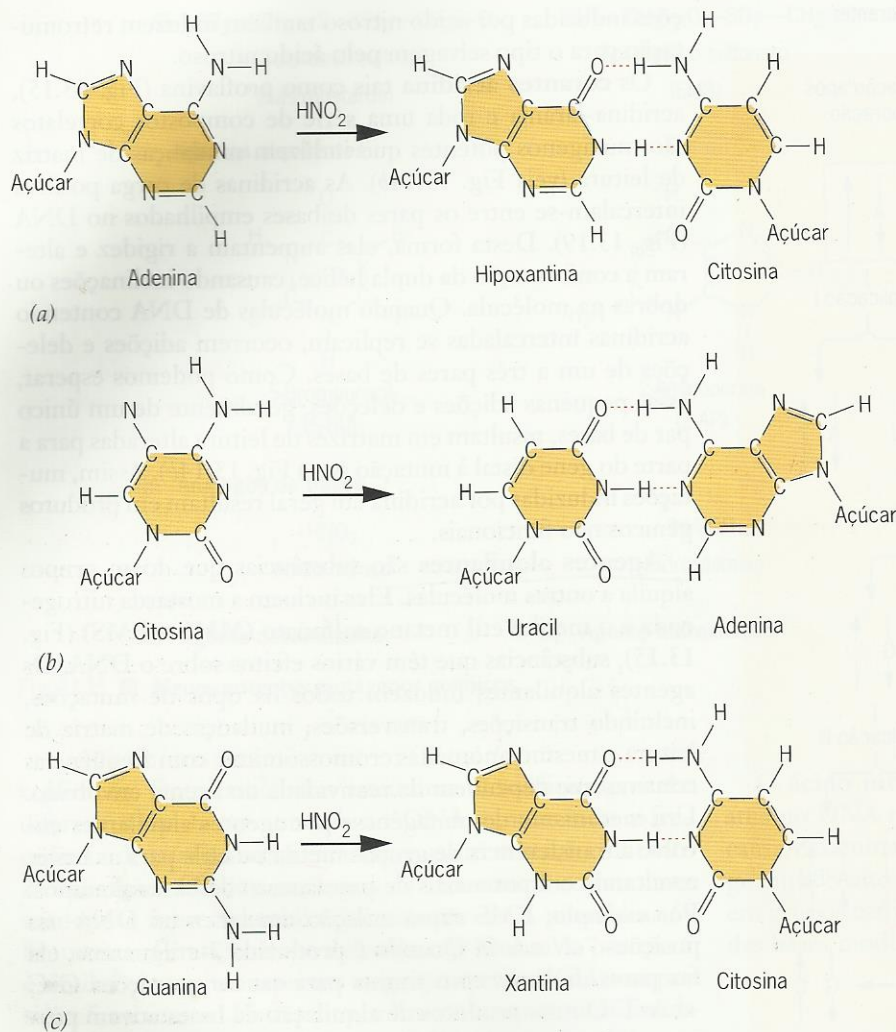


Fig. 13.18 ■ O ácido nitroso induz mutações por desaminação oxidativa das bases no DNA. O ácido nitroso converte (a) adenina em hipoxantina, causando transições A-T → G-C; (b) citosina em uracil, causando transições G-C → A-T; e (c) guanina em xantina, que não é mutagênica. Juntos, os efeitos do ácido nitroso sobre adenina e citosina explicam sua capacidade de induzir transições em ambos os sentidos, A-T ↔ G-C.

hidroxilamina. (2) Aquelas com um par de bases G-C no sítio mutante serão induzidas a reverter por hidroxilamina. Assim, a hidroxilamina pode ser usada para determinar se uma mutação em particular era uma transição A-T → G-C ou G-C → A-T.

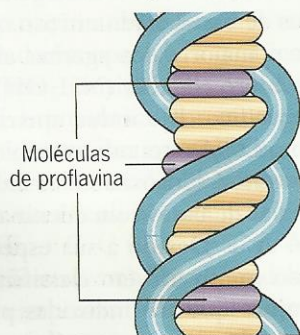


Fig. 13.19 ■ Intercalação de proflavina na dupla hélice de DNA. Estudos de difração de raios X mostraram que estes corantes de acridina de carga positiva ficam inseridos entre os pares de bases empilhados.

Mutações Induzidas por Radiação

Uma parte do espectro eletromagnético (Fig. 13.20) com comprimentos de onda menores e energia mais alta que a luz visível é subdividida em **radiação ionizante** (raios X, raios gama e raios cósmicos) e **radiação não ionizante** (luz ultravioleta). As radiações ionizantes são de alta energia e são úteis para diagnóstico médico porque penetram nos tecidos vivos a distâncias substanciais. No processo, esses raios de alta energia colidem com átomos e causam a liberação de elétrons, criando radicais livres ou íons de carga positiva. Os íons, por sua vez, colidem com outras moléculas e causam a liberação de elétrons adicionais. O resultado é que se forma um cone de íons ao longo da via de cada raio de alta energia à medida que ele passa através dos tecidos. Este processo de ionização é induzido por raios X originados em máquinas, por prótons e nêutrons bem como pelos raios alfa, beta e gama liberados por isótopos radioativos tais como ^{32}P , ^{35}S e o urânio-238 usado em reatores nucleares.

Os raios ultravioleta, tendo menos energia que as radiações ionizantes, penetram apenas na camada superficial de células em plantas e animais superiores e não causam io-

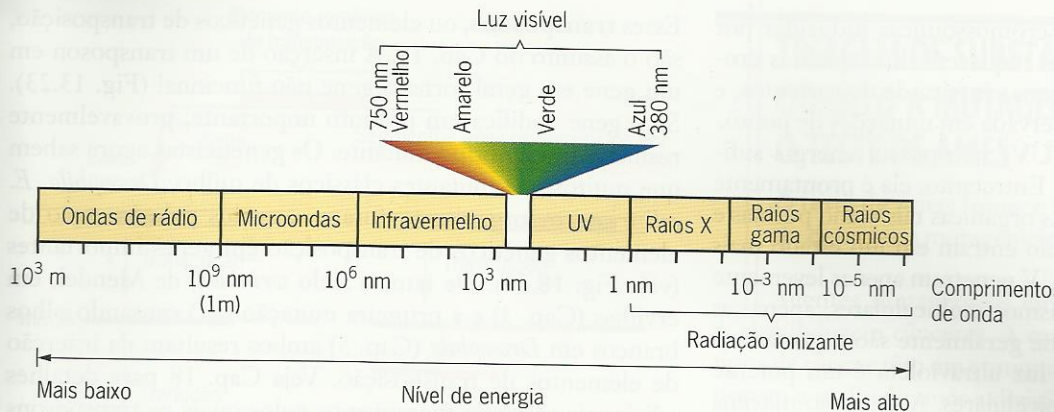


Fig. 13.20 ■ O espectro eletromagnético.

nização. Os raios ultravioleta dissipam sua energia para os átomos que eles encontram, elevando os elétrons nas órbitas mais externas para níveis energéticos mais altos, um estado chamado **excitação**. As moléculas contendo átomos em formas iônicas ou estados excitados são quimicamente mais reativas que aquelas contendo átomos em seus estados normais estáveis. A reatividade aumentada dos átomos presentes em moléculas de DNA é responsável pela maioria da mutagenicidade da radiação ionizante e da luz ultravioleta.

Os raios X e outras formas de radiação ionizante são quantificados em unidades **roentgen (r)**, que são medidas do número de ionizações por unidade de volume em um conjunto padrão de condições. Especificamente, uma unidade roentgen é uma quantidade de radiação ionizante que produz $2,083 \times 10^9$ pares de íons em um centímetro cúbico de ar a 0°C e em pressão de 760 mm de mercúrio. Note que a dose de radiação em unidades roentgen não envolve uma escala de tempo. A mesma dose pode ser obtida por baixa intensidade de irradiação durante longo período de tempo ou alta intensidade de irradiação por curto período de tempo. Este ponto é importante porque na maioria dos estudos a

freqüência de mutações de ponto induzidas é diretamente proporcional à dose de irradiação (Fig. 13.21). Por exemplo, raios X em espermatozoides de *Drosophila* causam aumento de aproximadamente 3% na taxa de mutação para cada 1.000 r de aumento na dose de irradiação. Esta correlação linear mostra que a indução de mutações por raios X exibe uma cinética de *single-hit*, o que significa que cada mutação resulta de um único evento de ionização. Isto é, cada ionização tem uma probabilidade fixa de induzir uma mutação em um conjunto padrão de condições.

Qual o nível seguro de irradiação? O desenvolvimento e o uso da bomba atômica e os acidentes em usinas nucleares criaram uma preocupação sobre exposição a radiações ionizantes. A correlação linear entre taxa de mutação e dose de irradiação indica que não existe nível seguro de irradiação. Os resultados indicam que, quanto maior a dose de irradiação, maior a taxa de mutação, e quanto menor a dose, menor a taxa de mutação. Mesmo níveis muito baixos de irradiação têm alguma probabilidade, baixa mas real, de induzir mutações.

Em espermatozoides de *Drosophila*, a irradiação crônica (baixos níveis de irradiação por longos períodos de tempo) é tão eficaz em induzir mutações quanto a irradiação aguda (a mesma dose total de irradiação ministrada em alta intensidade por curtos períodos de tempo). Entretanto, em camundongos, a irradiação crônica resulta em menos mutações que a mesma dose de irradiação aguda. Além disso, quando camundongos são tratados com doses intermitentes de irradiação, a freqüência de mutação é levemente menor do que quando são tratados com a mesma quantidade total de irradiação em uma dose contínua. A resposta diferencial de moscas-das-frutas e mamíferos a irradiação crônica parece resultar de diferenças na eficiência com que estas espécies reparam danos ao DNA induzidos por irradiação. Os mecanismos de reparo podem existir em espermatogônias e ovócitos de mamíferos mas que não funcionam em espermatozoides de *Drosophila*. Entretanto, devemos destacar que todos estes tratamentos de irradiação são mutagênicos, embora em graus diferentes, tanto em *Drosophila* quanto em mamíferos.

A radiação ionizante também induz grandes mudanças na estrutura do cromossomo, incluindo deleções, duplicações, inversões e translocações (Cap. 6). Essas anomalias cromos-

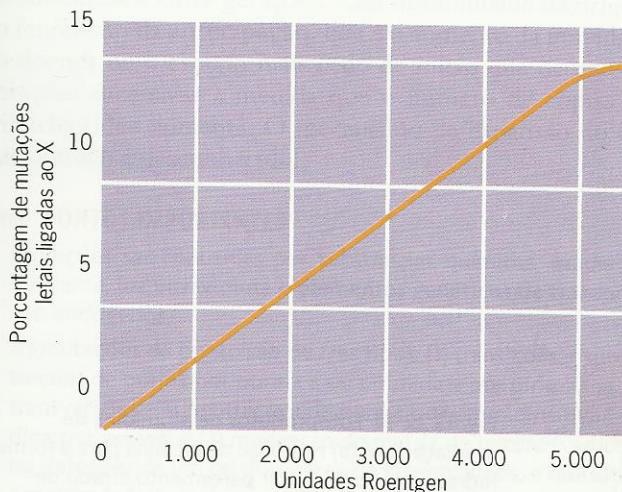


Fig. 13.21 ■ Correlações entre dose de irradiação e freqüência de mutação em *Drosophila*.

sômicas resultam de quebras cromossômicas induzidas por radiação. Como tais anomalias requerem duas quebras cromossômicas, elas apresentam uma cinética de dois eventos, e não de um evento, como a observada em mutações de ponto.

A **radiação ultravioleta (UV)** não possui energia suficiente para induzir ionizações. Entretanto, ela é prontamente absorvida por muitas moléculas orgânicas tais como purinas e pirimidinas no DNA, que então entram em um estado mais reativo, ou excitado. Os raios UV penetram apenas levemente nos tecidos. Assim, em organismos multicelulares, apenas as camadas de células da epiderme geralmente são expostas aos efeitos da UV. Entretanto, a luz ultravioleta é um potente mutágeno para organismos unicelulares. A absorção máxima de UV pelo DNA está em um comprimento de onda de 254 nm. Mutagenicidade máxima também ocorre em 254 nm, sugerindo que o processo de mutação induzido por UV é diretamente mediado pela absorção de UV por parte de purinas e pirimidinas. Estudos *in vitro* mostram que as pirimidinas absorvem fortemente a 254 nm e, como resultado, tornam-se muito reativas. Dois produtos importantes da absorção de UV pelas pirimidinas (timina e citosina) são os hidratos de pirimidina e os dímeros de pirimidina (Fig. 13.22). Os dímeros de timina causam mutações de dois modos. (1) Os dímeros perturbam a estrutura das duplas hélices de DNA e interferem na replicação precisa do DNA. (2) Ocorrem erros durante os processos celulares que reparam defeitos no DNA, tais como dímeros de timina induzidos por UV (veja a seção Mecanismos de Reparo do DNA, mais adiante neste capítulo).

Mutações Induzidas por Elementos Genéticos de Transposição

Os organismos vivos contêm marcantes elementos de DNA que podem mover-se de um local no genoma para outro.

Esses **transposons**, ou elementos genéticos de transposição, são o assunto do Cap. 18. A inserção de um transposon em um gene em geral torna o gene não funcional (Fig. 13.23). Se o gene codifica um produto importante, provavelmente resultará um fenótipo mutante. Os geneticistas agora sabem que muitos dos mutantes clássicos de milho, *Drosophila*, *E. coli* e outros organismos foram causados pela inserção de elementos genéticos de transposição em genes importantes (veja Fig. 18.21). De fato, o alelo *wrinkled*, de Mendel, em ervilhas (Cap. 3) e a primeira mutação (*w¹*) causando olhos brancos em *Drosophila* (Cap. 5) ambos resultam da inserção de elementos de transposição. Veja Cap. 18 para detalhes adicionais sobre os mecanismos pelos quais os transposons se movem e, no processo, produzem mutações.

Repetições Expandidas de Trinucleotídeos e Doenças Humanas Herdadas

Todos os tipos de mutações discutidas nas seções anteriores deste capítulo ocorrem em humanos. Além disso, ocorre outro tipo de mutação que está associada a doenças humanas. As seqüências repetidas de um a seis pares de nucleotídeos são conhecidas como **repetições simples em tandem**. Tais repetições estão dispersas pelo genoma humano. As repetições de três pares de nucleotídeos, **repetições de trinucleotídeos**, podem aumentar em número de cópias e causar doenças herdadas em humanos. Vários trinucleotídeos comprovadamente sofrem tais aumentos no número de cópias. As repetições expandidas do trinucleotídeo CGG no sítio *FRAXA* do cromossomo X são responsáveis pela síndrome do X frágil, a forma mais comum de retardo mental herdado em humanos (Cap. 5). Cromossomos X normais contêm de 6 a cerca de 50 cópias da repetição CGG no sítio *FRAXA*.

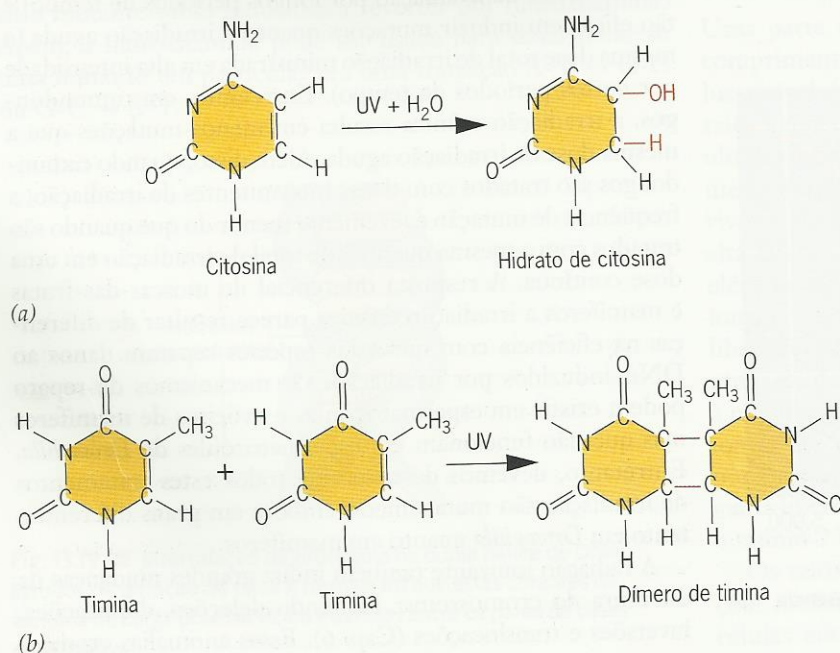


Fig. 13.22 ■ Fotoprodutos de pirimidina de irradiação UV. (a) Hidrólise de citosina para a forma hidrato que pode causar pareamento errado de bases durante a replicação. (b) Ligação cruzada de moléculas adjacentes de timina para formar dímeros de timina, que bloqueiam a replicação do DNA.

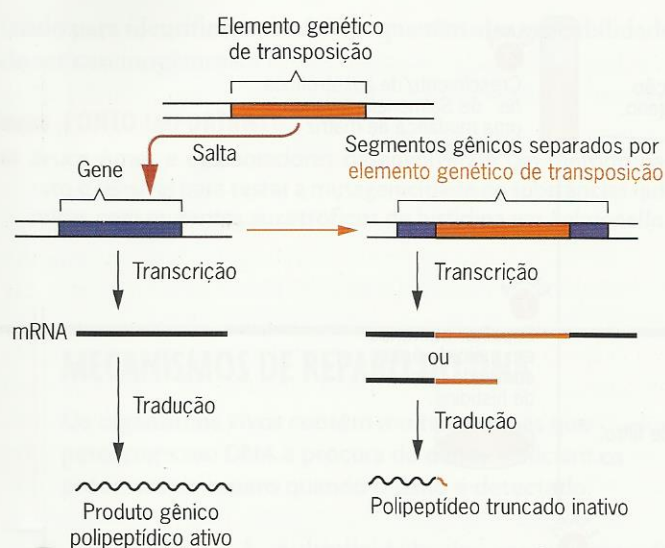


Fig. 13.23 ■ Mecanismo de mutação induzida por transposon. A inserção de um elemento genético de transposição (vermelho) em um gene tipo selvagem (à esquerda) geralmente torna o gene não funcional (à direita). Um produto gênico truncado em geral resulta de sinais de término de transcrição ou tradução, ou ambos, situados dentro do transposon.

Cromossomos X mutantes contêm até 1.000 cópias da repetição em tandem CGG neste sítio.

As repetições dos trinucleotídeos CAG e CTG estão envolvidas em várias doenças neurológicas herdadas, incluindo doença Huntington (discutida em detalhes no Cap. 17), distrofia miotônica, doença Kennedy, atrofia palidoluisiana dentatorrubral, doença Machado-Joseph e ataxia espinocerebelar. Em todos estes distúrbios neurológicos, a gravidade da doença está correlacionada com o número de cópias do trinucleotídeo – quanto maior o número de cópias, mais graves os sintomas da doença. Além disso, os trinucleotídeos expandidos associados a estas doenças são instáveis em células somáticas e entre gerações. Essa instabilidade dá origem ao fenômeno da **antecipação**, que é o aumento da gravidade da doença ou o estágio mais cedo de início que ocorre em gerações sucessivas à medida que o número de cópias do trinucleotídeo aumenta. O mecanismo de expansão do trinucleotídeo é desconhecido.

PONTOS IMPORTANTES

- Mutações são induzidas por substâncias químicas, irradiação ionizante, luz ultravioleta e elementos genéticos de transposição endógenos.
- As mutações de ponto são de três tipos: (1) transições – substituições de purina por purina e de pirimidina por pirimidina; (2) transversões – substituições de purina por pirimidina e pirimidina por purina; e (3) mudanças de matriz de leitura – adições ou deleções de um ou dois pares de nucleotídeos, o que altera a matriz de leitura do gene distal ao sítio da mutação.
- Várias doenças hereditárias humanas são causadas por repetições expandidas de trinucleotídeos.

TRIAGEM DE SUBSTÂNCIAS QUÍMICAS QUANTO A MUTAGENICIDADE: O TESTE DE AMES

O teste de Ames fornece um método simples e barato de detectar a mutagenicidade de substâncias químicas.

Agentes mutagênicos também são **carcinógenos**, isto é, eles induzem cânceres. A característica que as centenas de tipos de câncer têm em comum é que as células malignas continuam a se dividir após a divisão celular ter parado em células normais. Logicamente, a divisão celular, como todos os outros processos biológicos, está sob controle genético. Genes específicos codificam produtos que regulam a divisão celular em resposta a sinais intracelulares, intercelulares e ambientais. Quando tais genes passam por mutação para estados não funcionais, às vezes resulta uma divisão celular descontrolada. Claramente, desejamos evitar exposição a agentes mutagênicos e carcinogênicos. Entretanto, nossa sociedade tecnológica depende de uma ampla utilização de substâncias químicas tanto na indústria quanto na agricultura. Centenas de substâncias químicas novas são produzidas a cada ano, e a mutagenicidade e a carcinogenicidade destas substâncias precisam ser avaliadas antes que seu uso se torne generalizado.

Tradicionalmente, a carcinogenicidade de substâncias químicas tem sido testada em roedores, em geral camundongos neonatos. Estes estudos envolvem procedimentos de alimentação ou injeção com as substâncias que estão sendo testadas e subsequentemente exame dos animais quanto a tumores. Os testes de mutagenicidade são feitos de modo similar. Entretanto, como a mutação é um evento de baixa frequência e como a manutenção de grandes populações de roedores é muito cara, os testes têm sido relativamente insensíveis; isto é, baixos níveis de mutagenicidade podem não ser detectados.

Bruce Ames e colaboradores desenvolveram técnicas sensíveis que permitem testar rapidamente a mutagenicidade de grandes números de substâncias e a um custo relativamente baixo. Ames e colaboradores construíram linhagens auxotróficas da bactéria *Salmonella typhimurium* que leva vários tipos de mutações – transições, transversões e mudanças de matriz de leitura – em genes necessários para a biossíntese do aminoácido histidina. Eles monitoraram a reversão destes mutantes auxotróficos para prototróficos colocando um número conhecido de bactérias mutantes em um meio sem histidina e registrando o número de colônias produzidas por revertentes prototróficos. Como algumas substâncias são mutagênicas apenas para o DNA replicante, eles adicionaram uma pequena quantidade de histidina ao meio, o suficiente para permitir algumas divisões celulares mas não a formação de colônias. Eles dosaram a mutagenicidade de uma substância comparando a frequência de reversão em sua presença com a frequência de reversão espontânea (Fig. 13.24). Eles avaliaram sua capacidade de induzir tipos dife-

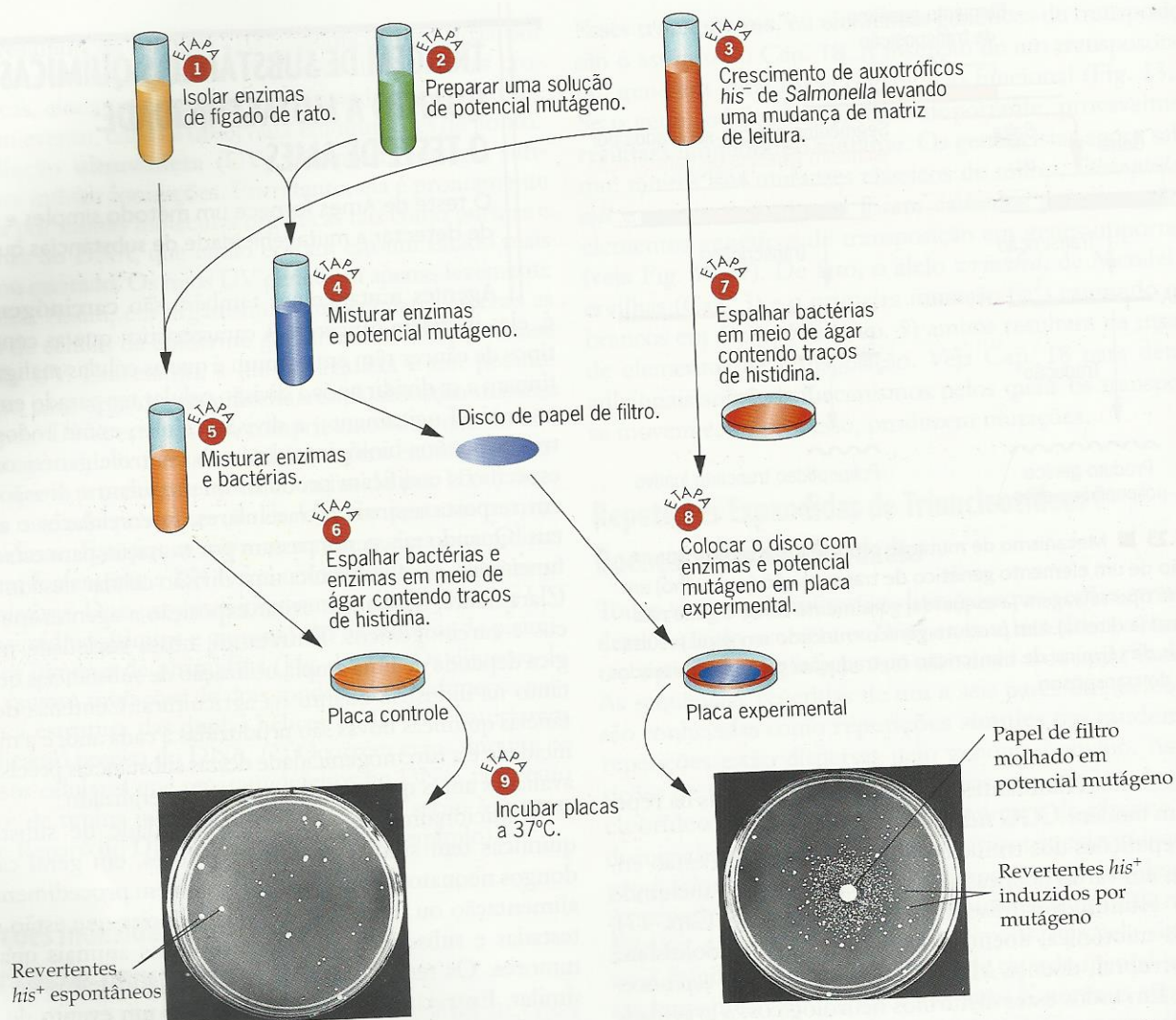


Fig. 13.24 ■ O teste de Ames para mutagenicidade. O meio em cada placa de Petri contém traços de histidina e um número conhecido de células *his⁻* de uma "linhagem testadora" específica de *Salmonella typhimurium* possuindo uma mudança de matriz de leitura. A placa de células *his⁻* de uma "linhagem testadora" específica de *Salmonella typhimurium* possuindo uma mudança de matriz de leitura. A placa experimental à direita mostra a frequência de reversão induzida pelo potencial mutágeno – neste caso, o carcinógeno 2-aminofluoreno.

rentes de mutações usando um conjunto de linhagens testadoras que levam tipos diferentes de mutações – uma linhagem com uma transição, uma com uma mudança de matriz de leitura e assim em diante.

Durante um período de vários anos durante o qual eles testaram milhares de substâncias químicas diferentes, Ames e seus colaboradores observaram uma correlação maior que 90% entre a mutagenicidade e a carcinogenicidade das substâncias testadas. Inicialmente, eles descobriram que vários carcinógenos potentes não são mutagênicos para as linhagens testadas. Subseqüentemente, eles descobriram que muitos destes carcinógenos são metabolizados em derivados fortemente mutagênicos em células eucarióticas. Assim, Ames e seus associados adicionaram um extrato de fígado de rato a seus sistemas de análise na tentativa de detectar a mutagenicidade de derivados metabólicos das substâncias

em teste. O acoplamento do sistema de ativação de fígado de rato aos testes de mutagenicidade microbiana ampliou consideravelmente a utilidade do sistema. Por exemplo, nitratos (encontrados em carnes cozidas) em si não são mutagênicos ou carcinogênicos. Entretanto, em células eucarióticas, os nitratos são convertidos em nitrosaminas, que são altamente mutagênicas e carcinogênicas. Os testes de mutagenicidade de Ames demonstraram a presença de mutágenos de matriz de leitura em vários componentes de fumo de cigarro condensado fracionado quimicamente. Em alguns casos, a ativação pela preparação de extrato de fígado foi necessária para a mutagenicidade. Em outros casos, a ativação não foi necessária. O teste de Ames fornece um procedimento rápido, barato e sensível para testar a mutagenicidade de substâncias químicas. Como as substâncias químicas mutagênicas também são carcinógenas, o teste de Ames pode ser

usado para identificar substâncias que têm alta probabilidade de ser carcinogênicas.

PONTO IMPORTANTE

■ Bruce Ames e colaboradores desenvolveram um método barato e sensível para testar a mutagenicidade de substâncias químicas com mutantes auxotróficos de histidina em *Salmonella*.

MECANISMOS DE REPARO DO DNA

Os organismos vivos contêm muitas enzimas que percorrem seu DNA à procura de danos e iniciam os processos de reparo quando o dano é detectado.



A multiplicidade dos mecanismos de reparo desenvolvidos em organismos que variam de bactérias a humanos documentam enfaticamente a importância de manter a mutação em um nível

tolerável. Por exemplo, as *E. coli* possuem cinco mecanismos bem caracterizados de reparo de defeitos no DNA: (1) reparo dependente de luz ou fotorreativação, (2) reparo por excisão, (3) reparo de pareamento errado, (4) reparo após replicação e (5) o sistema de reparo propenso a erro (resposta SOS). Além disso, existem pelo menos dois tipos diferentes de reparo por excisão, e as vias de reparo por excisão podem ser iniciadas por várias enzimas diferentes, cada uma agindo em um tipo específico de dano ao DNA. Os mamíferos parecem possuir todos os mecanismos de reparo encontrados em *E. coli* exceto a fotorreativação. Como a maioria das células de mamíferos não têm acesso a luz, a fotorreativação seria de valor relativamente pequeno para elas.

A importância das vias de reparo do DNA para a saúde humana está clara. Distúrbios herdados tais como o xeroderma pigmentoso, que foi discutido no começo deste capítulo, documentam as graves consequências de defeitos no reparo do DNA. Discutiremos alguns destes distúrbios herdados em uma seção subsequente deste capítulo.

Reparo Dependente de Luz

O reparo dependente de luz ou fotorreativação do DNA em bactérias é feito por uma enzima ativada por luz chamada **DNA-fotoliase**. Quando o DNA é exposto a luz ultravioleta, dímeros de timina são produzidos por ligações cruzadas covalentes entre timinas adjacentes (veja Fig. 13.22). A DNA-fotoliase reconhece os dímeros de timina no DNA e se liga a eles, usando a energia da luz para cortar as ligações cruzadas covalentes (Fig. 13.25). A fotoliase liga-se aos dímeros de timina no DNA no escuro, mas não pode catalisar a clivagem destas ligações juntando as timinas sem a energia derivada da luz visível, especificamente a luz da região azul do espectro. A fotoliase também corta dímeros de citosina e

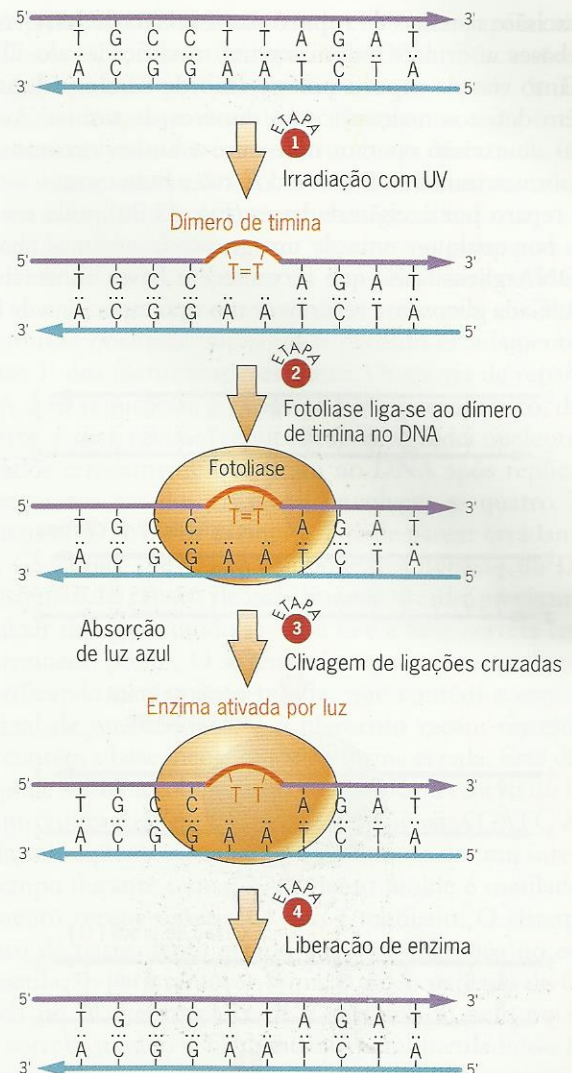


Fig. 13.25 ■ Clivagem de ligações cruzadas de dímeros de timina por fotoliase ativada pela luz. As setas indicam a polaridade oposta dos filamentos complementares do DNA.

dímeros de citosina-timina. Assim, quando a luz ultravioleta é usada para induzir mutações em bactérias, as células irradiadas são cultivadas no escuro por algumas gerações para maximizar a frequência de mutação.

Reparo por Excisão

O reparo por excisão de DNA danificado envolve pelo menos três etapas. Na etapa 1, uma endonuclease de reparo do DNA ou complexo enzimático contendo endonuclease reconhece a base (ou bases) danificada no DNA, liga-se a ela e a excisa. Na etapa 2, uma DNA-pólimerase preenche o espaço usando o filamento complementar não danificado de DNA como molde. Na etapa 3, a enzima DNA-ligase fecha a falha deixada pela DNA-pólimerase para completar o processo de reparo. Existem dois tipos principais de reparo

por excisão: sistemas de **reparo por excisão de base** removem bases anormais quimicamente modificadas do DNA, enquanto vias de **reparo por excisão de nucleotídeos** removem defeitos maiores como dímeros de timina. Ambas as vias de excisão operam no escuro e ambas ocorrem por mecanismos muito similares em *E. coli* e humanos.

O reparo por excisão de bases (Fig. 13.26) pode ser iniciado por qualquer uma de um grupo de enzimas chamadas DNA-glicosilases, que reconhecem bases anormais no DNA. Cada glicosilase reconhece um tipo específico de base

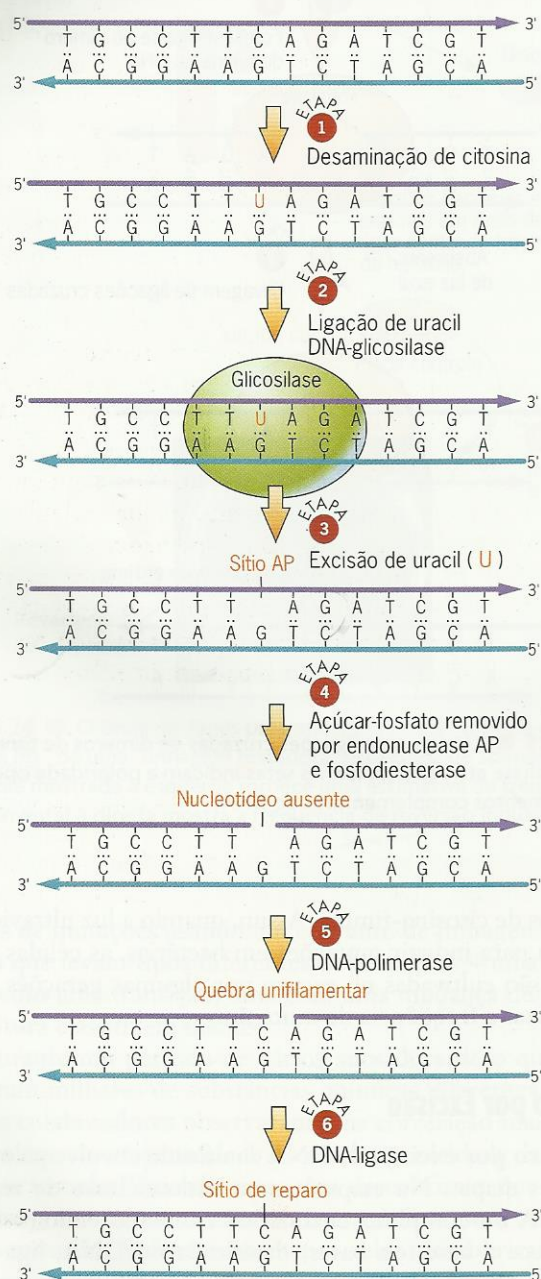


Fig. 13.26 ■ Reparo do DNA pela via de excisão de bases. O reparo por excisão de bases pode ser iniciado por qualquer uma das diferentes DNA-glicosilases. Nos exemplos mostrados, a uracil DNA-glicosilase começa o processo de reparo.

alterada, tal como bases desaminadas, bases oxidadas e assim em diante (etapa 2). As glicosilases clivam a ligação glicosídica entre a base anormal e 2-desoxirribose, criando sítios apurínicos ou apirimidínicos (sítios AP) com bases ausentes (etapa 3). Os sítios AP são reconhecidos por enzimas chamadas endonucleases AP, que atuam juntas com fosfodiesterases para excisar os grupos açúcar-fosfato nesses sítios (etapa 4). A DNA-polimerase então substitui o nucleotídeo ausente de acordo com especificações do filamento complementar (etapa 5), e a DNA-ligase fecha o corte (etapa 6).

O reparo por excisão de nucleotídeos remove lesões maiores, como os dímeros de timina e bases com grupos laterais grandes do DNA. No reparo por excisão de nucleotídeos, uma única atividade de excisão de nucleases produz cortes em ambos os lados do(s) nucleotídeo(s) danificado(s) e excisa um oligonucleotídeo contendo a(s) base(s) danificada(s). Esta nuclease é chamada **excinuclease** para distingui-la das endonucleases e exonucleases que têm outros papéis no metabolismo do DNA.

A via de reparo por excisão de nucleotídeo de *E. coli* é mostrada na Fig. 13.27. Em *E. coli*, a atividade de excinuclease requer os produtos de três genes, *uvrA*, *uvrB* e *uvrC* (chamados *uvr* por reparo de UV). Uma proteína trimérica contendo dois polipeptídeos UvrA e um UvrB reconhece o defeito no DNA, liga-se a ele e usa a energia do ATP para dobrar o DNA no sítio danificado. O dímero UvrA é então liberado, e a proteína UvrC liga-se ao complexo UvrB/DNA. A proteína UvrB corta a quinta ligação fosfodiéster do(s) nucleotídeo(s) danificado(s) no lado 3', e a proteína UvrC hidrolisa a oitava ligação fosfodiéster a partir do dano no lado 5'. O produto do gene *uvrD*, a DNA-helicase II, libera o dodecâmero excisado. Nas últimas duas etapas da via, a DNA-polimerase I preenche o espaço, e a DNA-ligase fecha o corte restante na molécula de DNA.

O reparo por excisão de nucleotídeos em humanos ocorre por uma via similar à de *E. coli*, mas envolve cerca de quatro vezes mais proteínas. Em humanos, a atividade de excinuclease contém 15 polipeptídeos. A proteína XPA (de proteína A de xeroderma pigmentoso) reconhece o(s) nucleotídeo(s) danificado(s) e se liga a ele(s) no DNA. Ela então recruta outras proteínas necessárias para atividade de excinuclease. Em humanos, o oligômero excisado tem 24 a 32 nucleotídeos de tamanho, em vez dos 12 removidos em *E. coli*. O espaço é preenchido ou pela DNA-polimerase δ ou pela ϵ em humanos, e a DNA-ligase completa o trabalho.

Outros Mecanismos de Reparo do DNA

Durante os últimos anos, pesquisas sobre os mecanismos de reparo do DNA demonstraram a presença de um grupo de enzimas de reparo do DNA que constantemente percorrem o DNA em busca de danos, variando desde a presença de dímeros de timina induzidos por luz ultravioleta até modificações muito diversas e numerosas para serem descritas aqui. Novos resultados deste trabalho mostram que várias

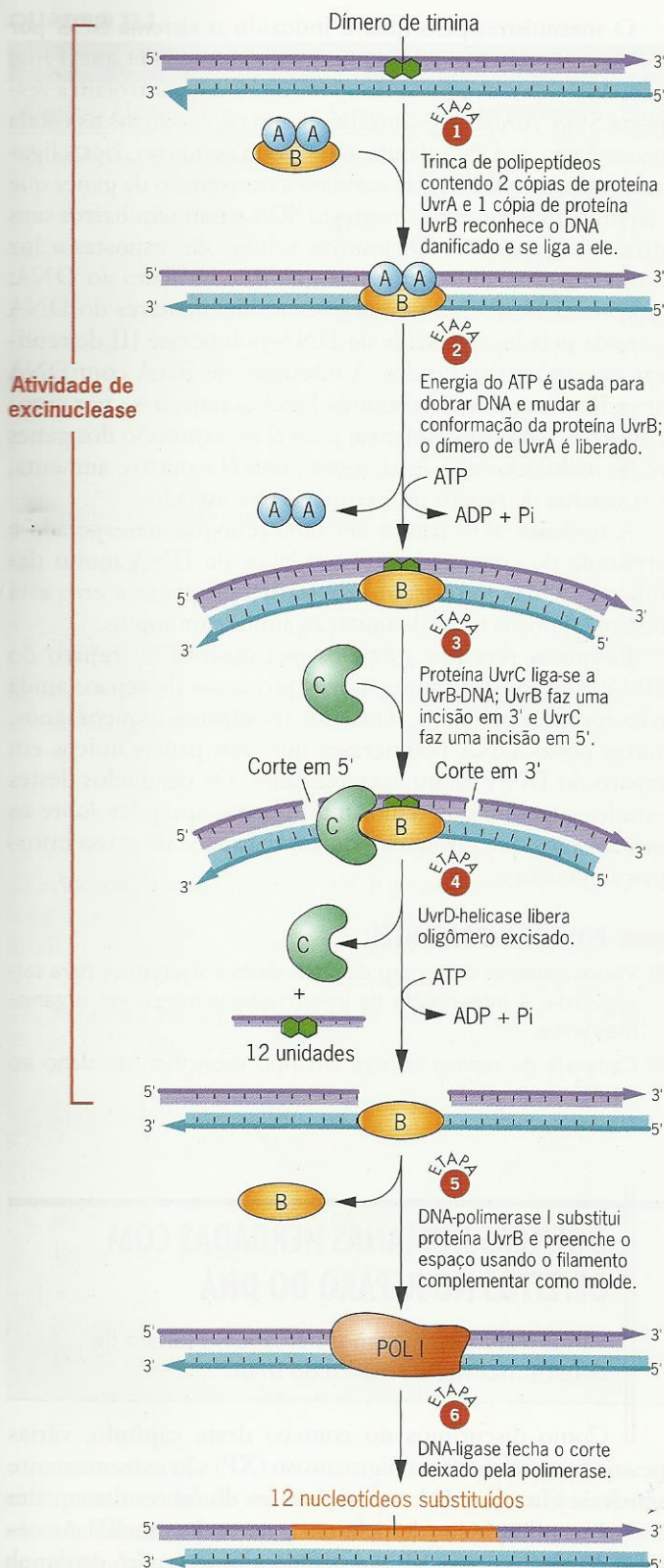


Fig. 13.27 ■ Reparo do DNA pela via de excisão de nucleotídeos em *E. coli*. A atividade de excinuclease (nuclease de excisão) requer os produtos de três genes, *uvrA*, *uvrB* e *uvrC*. A excisão de nucleotídeos ocorre por uma via similar em humanos, exceto por estarem envolvidas muito mais proteínas e um oligômero com 24 a 32 nucleotídeos ser excisado.

DNA-polimerases antes não conhecidas têm papéis críticos em diversos processos de reparo do DNA. Discussões detalhadas destes importantes processos de reparo do DNA estão além do escopo deste texto. Entretanto, a importância destes mecanismos de reparo não pode ser descartada. O que é mais importante para a sobrevivência de uma espécie do que manter a integridade de sua estrutura genética?

No Cap. 10, examinamos o mecanismo pelo qual a atividade de exonuclease 3' → 5' das DNA-polimerases faz uma revisão nos filamentos de DNA durante sua síntese, removendo qualquer nucleotídeo pareado erradamente nas pontas 3' dos filamentos crescentes. Outra via de reparo do DNA após replicação, o **reparo de malpareamento**, dá um suporte a esta revisão replicativa corrigindo nucleotídeos pareados erradamente que ficam no DNA após replicação. Pareamentos errados em geral envolvem as quatro bases normais no DNA. Por exemplo, T pode parrear erradamente com G. Como T e G são componentes normais do DNA, os sistemas de reparo de pareamento errado precisam determinar de algum modo se T ou G é a base correta em um determinado ponto. O sistema de reparo faz esta distinção identificando o filamento molde, que contém a sequência original de nucleotídeos, e o filamento recém-sintetizado, que contém a base incorporada de forma errada. Esta distinção pode ser feita com base no padrão de metilação do DNA recém-replicado. Em *E. coli*, A nas sequências GATC é metilada subsequente a sua síntese. Assim, ocorre um intervalo de tempo durante o qual o filamento molde é metilado e o filamento recém-sintetizado não é metilado. O sistema de reparo de pareamento errado usa esta diferença no estado de metilação para remover o nucleotídeo pareado de forma errada no filamento nascente e substituí-lo pelo nucleotídeo correto usando o filamento parental metilado do DNA como molde.

Em *E. coli*, o reparo de malpareamento requer os produtos de quatro genes, *mutH*, *mutL*, *mutS* e *mutU* (= *uvrD*). A proteína MutS reconhece pareamentos errados e se liga a eles para iniciar o processo de reparo. As proteínas MutH e MutL então se unem ao complexo. MutH contém uma **atividade de endonuclease específica de GATC** que corta o filamento não metilado em sítios GATC hemimetilados (isto é, metade metilados) em 5' ou em 3' do pareamento errado. Os sítios de incisão podem estar a 1.000 pares de exonucleotídeos ou mais do malpareamento. O processo subsequente de excisão requer MutS, MutL, DNA-helicase II (MutU) e uma exonuclease apropriada. Se a incisão ocorre em uma sequência GATC a 5' do malpareamento, é necessária uma exonuclease 5' → 3' como a exonuclease VII de *E. coli*. Se a incisão ocorre a 3' do malpareamento, é necessária uma atividade 3' → 5' de atividade exonuclease como a da exonuclease I de *E. coli*. Após o processo de excisão ter removido o nucleotídeo malpareado do filamento não metilado, a DNA-polimerase III preenche o espaço, de até 1.000 pb, e a DNA-ligase fecha o corte.

Foram identificados homólogos das proteínas MutS e MutL de *E. coli* em fungos, plantas e mamíferos, uma indica-

ção de que vias similares de reparo de malpareamento ocorrem em eucariontes. De fato, a excisão de malpareamento foi demonstrada *in vitro* com extratos nucleares preparados de células humanas. Assim, o reparo de malpareamento é provavelmente um mecanismo universal ou quase universal para resguardar a integridade da informação genética estocada na dupla hélice de DNA.

Em *E. coli*, reparo dependente de luz, reparo de excisão e reparo de malpareamento podem ser eliminados por mutações nos genes *phr* (fotorreativação), *uvr* e *mut*, respectivamente. Em mutantes deficientes de mais de um destes mecanismos de reparo, um outro sistema de reparo do DNA, chamado **reparo de pós-replicação**, funciona. Quando a DNA-polimerase III encontra um dímero de timina em um filamento molde, seu progresso é bloqueado. A DNA-polimerase reinicia a síntese de DNA em alguma posição após o dímero, deixando um espaço no filamento nascente oposto ao dímero no filamento molde. Neste ponto, a sequência original de nucleotídeos foi perdida de ambos os filamentos da dupla hélice prole. A molécula de DNA danificada é reparada por um processo de reparo dependente de recombinação mediado pelo produto do gene *recA* de *E. coli*. A proteína RecA, que é necessária para recombinação homóloga, estimula a troca de filamentos isolados entre as duplas hélices homólogas. Durante o reparo de pós-replicação, a proteína RecA liga-se ao filamento único de DNA no espaço e medeia o pareamento com o segmento homólogo da dupla hélice irmã. O espaço em oposição ao dímero é preenchido com o filamento homólogo de DNA da molécula irmã de DNA. O espaço resultante na dupla hélice irmã é preenchido pela DNA-polimerase, e o corte é fechado pela DNA-ligase. O dímero de timina permanece no filamento molde da prole original da molécula de DNA, mas o filamento complementar agora está intacto. Se o dímero de timina não é removido pelo sistema de reparo de excisão de nucleotídeos, esse reparo após a replicação deve ser repetido depois de cada rodada de replicação do DNA.

Os sistemas de reparo do DNA descritos até agora são bem precisos. Entretanto, quando o DNA de *E. coli* é muito danificado por agentes mutagênicos tais como luz UV, as células adotam algumas etapas drásticas na tentativa de sobreviver. Elas passam por um sistema chamado **resposta SOS**, durante a qual é produzida toda uma bateria de reparo do DNA, recombinação e proteínas de replicação. Duas dessas proteínas, codificadas pelos genes *umuC* e *umuD* (mutáveis por UV), são subunidades da DNA-polimerase V, uma enzima que catalisa a replicação do DNA em regiões danificadas do cromossomo – regiões onde a replicação pela DNA-polimerase III está bloqueada. A DNA-polimerase V permite que a replicação continue através de segmentos danificados dos filamentos moldes, muito embora a sequência de nucleotídeos na região danificada não possa ser replicada com precisão. Este sistema de **reparo propenso a erro** elimina espaços nos filamentos recém-sintetizados em oposição a nucleotídeos danificados nos filamentos moldes, mas, ao fazer isto, aumenta a frequência de erros de replicação.

O mecanismo pelo qual é induzido o sistema SOS por dano ao DNA foi esclarecido em considerável detalhe. Duas proteínas chaves reguladoras, LexA e RecA, controlam a resposta SOS. Ambas são sintetizadas em níveis baixos na célula na ausência de DNA danificado. Nesta condição, LexA liga-se a regiões do DNA que regulam a transcrição de genes que são induzidos durante a resposta SOS e mantém baixos seus níveis de expressão. Quando as células são expostas a luz ultravioleta ou outros agentes que causam dano ao DNA, a proteína RecA liga-se a regiões unifilamentares do DNA causada pela incapacidade da DNA-polimerase III de replicar as regiões danificadas. A interação de RecA com DNA ativa RecA, que então estimula LexA a inativar-se por autoclivagem. Com LexA inativa, o nível de expressão dos genes SOS, incluindo *recA*, *lexA*, *umuC*, *umuD* e outros, aumenta, e o sistema de reparo propenso a erro é ativado.

A resposta SOS parece ser uma tentativa desesperada e arriscada de escapar dos efeitos letais de DNA muito danificado. Quando o sistema de reparo propenso a erro está funcionando, as taxas de mutação aumentam muito.

Pesquisas recentes sobre os mecanismos de reparo do DNA indicam que muitos novos processos de reparo ainda não foram elucidados. Durante os últimos poucos anos, várias novas DNA-polimerases que têm papéis únicos em reparo do DNA foram caracterizadas. Os resultados destes estudos sugerem que temos muito o que aprender sobre os mecanismos que salvaguardam a integridade de nossa informação genética.

■ PONTOS IMPORTANTES

- Vários sistemas de reparo do DNA desenvolveram-se para salvaguardar a integridade da informação genética em organismos vivos.
- Cada via de reparo corrige um tipo específico de dano ao DNA.

DOENÇAS HUMANAS HERDADAS COM DEFEITOS NO REPARO DO DNA

Vários distúrbios humanos herdados resultam de defeitos nas vias de reparo do DNA.

Como discutimos no começo deste capítulo, várias pessoas com xeroderma pigmentoso (XP) são extremamente sensíveis à luz do sol. A exposição à luz do sol resulta em alta frequência de câncer de pele em pacientes com XP. As células de pessoas com XP são deficientes no reparo de danos ao DNA induzidos por UV, tais como dímeros de timina. A síndrome XP pode resultar de defeitos em qualquer um de pelo menos oito genes diferentes. Os produtos de sete destes genes, *XPA*, *XPB*, *XPC*, *XPD*, *XPE*, *XPF* e *XPG*, são necessários para reparo por excisão de nucleotídeos (Quadro 13.1). Eles foram purificados e demonstrados como essen-

QUADRO 13.1

Doenças Humanas Hereditárias Causadas por Defeitos no Reparo do DNA

Doença Hereditária	Gene	Cromossomo	Função do Produto	Sintomas Principais
1. Xeroderma pigmentoso	<i>XPA</i>	9	Proteína de reconhecimento de dano ao DNA	Sensibilidade a UV, cânceres de pele de início cedo, distúrbios neurológicos
	<i>XPB</i>	2	Helicase 3' → 5'	
	<i>XPC</i>	3	Proteína de reconhecimento de dano ao DNA	
	<i>XPD</i>	19	Helicase 5' → 3'	
	<i>XPE</i>	11	Proteína de reconhecimento de dano ao DNA	
	<i>XPF</i>	16	Nuclease, incisão em 3'	
	<i>XPG</i>	13	Nuclease, incisão em 5'	
	<i>XPV</i>	6	DNA-polimerase de translesão	
2. Síndrome Cockayne	<i>CSA</i>	5	Proteína de reparo do DNA por excisão	Sensibilidade a UV, distúrbios neurológicos e do desenvolvimento, envelhecimento prematuro
	<i>CSB</i>	10	Proteína de reparo do DNA por excisão	
3. Tricotodistrofia	<i>TTDA</i>	6	Fator de transcrição basal IIIH	Sensibilidade a UV, distúrbios neurológicos, retardo mental
4. Ataxia-telangiectasia	<i>ATM</i>	11	Serina/treonina-cinase	Sensibilidade a radiação, instabilidade cromossômica, neurodegeneração progressiva de início cedo, propensão a câncer
5. Anemia Fanconi	<i>FA</i> (8 genes <i>A-H</i> , em 5 cromossomos diferentes)			Sensibilidade a agentes de ligação cruzada no DNA, instabilidade cromossômica, propensão a câncer
6. Síndrome Bloom	<i>BLM</i>	15	BLM RecQ-helicase	Instabilidade cromossômica, retardo mental, propensão a câncer
7. Síndrome Werner	<i>WRN</i>	8	WRN RecQ-helicase	Instabilidade cromossômica, neurodegeneração progressiva, propensão a câncer
8. Síndrome Rothmund-Thomson	<i>RECQL4</i>	8	RecQ-helicase L4	Instabilidade cromossômica, retardo mental, propensão a câncer
9. Síndrome Nijmegen por quebra	<i>NBSI</i>	8	Proteína de reconhecimento de quebra bifilamentar do DNA	Instabilidade cromossômica, microcefalia (crânio pequeno), propensão a câncer

ciais para a atividade de excinuclease. Como a atividade de excinuclease em humanos requer 15 polipeptídeos, a lista de genes *XP* provavelmente irá aumentar no futuro. Dois outros distúrbios humanos, a síndrome Cockayne e a tricotodistrofia, também resultam de defeitos no reparo de excisão de nucleotídeos. Pessoas com síndrome Cockayne exibem retardo no crescimento e nas habilidades mentais, mas não taxas aumentadas de câncer de pele. Pacientes com tricotodistrofia têm baixa estatura, cabelos quebradiços e

pele escamosa. Eles também têm habilidades mentais subdesenvolvidas. Os indivíduos com síndrome Cockayne ou com tricotodistrofia têm um defeito em um tipo de reparo por excisão que é acoplado a transcrição. Entretanto, os detalhes deste processo de reparo acoplado a transcrição ainda estão sendo desenvolvidos.

Além do dano às células da pele, algumas pessoas com *XP* desenvolvem anomalias neurológicas, que parecem resultar da morte prematura de células nervosas. Este efeito em célu-

las nervosas de longa vida pode ter implicações interessantes com relação às causas do envelhecimento. Uma teoria é que o envelhecimento resulta do acúmulo de mutações somáticas. Sendo assim, seria esperado que um sistema de reparo defeituoso acelere o processo de envelhecimento, e isto parece ser o caso em relação às células nervosas dos pacientes com XP. Entretanto, no momento, existem poucas evidências ligando mutação somática a senescência.

Ataxia-telangiectasia, anemia Fanconi, síndrome Bloom, síndrome Werner, síndrome Rothmund-Thomson e síndrome de quebras Nijmegen são outras seis doenças herdadas em humanos associadas a defeitos conhecidos no metabolismo do DNA. Todos os seis distúrbios exibem padrões autossômicos recessivos de herança e todos resultam em alto risco de malignidade, especialmente leucemia no caso da ataxia-telangiectasia e anemia Fanconi. As células dos pacientes com ataxia-telangiectasia exibem sensibilidade anormal a radiação ionizante, sugerindo um defeito no reparo de dano ao DNA induzido por radiação. As células das pessoas com anemia Fanconi são prejudicadas na remoção de ligações cruzadas interfilamentares de DNA, tais como as formadas pelo antibiótico mitomicina C. Pessoas com síndrome Bloom e síndrome de quebra Nijmegen exibem alta frequência de quebras cromossômicas que resultam em anomalias cromossômicas (Cap. 6) e trocas entre cromátides irmãs. A ataxia-telangiectasia é causada por defeitos em uma cinase envolvida no controle do ciclo celular, e a síndrome Bloom, a síndrome Werner e a síndrome Rothmund-Thomson resultam de alterações em DNA helicases específicas (membros da família de helicases RecQ). O Quadro 13.1 cita algumas das doenças humanas mais bem conhecidas que resultam de defeitos herdados nas vias de reparo do DNA. A demonstração de que o início do câncer colorretal não polipóide hereditário está associado a um defeito no reparo de pares de bases malpareados no DNA indica que defeitos similares em outras vias de reparo do DNA podem estar envolvidos no desenvolvimento de tipos específicos de câncer humano (Cap. 24).

■ PONTOS IMPORTANTES

- A importância das vias de reparo do DNA é documentada convincentemente pelos distúrbios hereditários humanos que resultam de defeitos no reparo do DNA.
- Alguns tipos de câncer também estão associados a defeitos nas vias de reparo do DNA.

MECANISMOS DE RECOMBINAÇÃO DO DNA

A recombinação entre moléculas homólogas de DNA envolve a atividade de várias enzimas que clivam, deselicoïdizam, estimulam invasões unifilamentares de duplas hélices, reparam e unem filamentos de DNA.

No Cap. 7, discutimos as principais características da recombinação entre cromossomos homólogos, mas não

consideramos os detalhes moleculares do processo. Como muitos dos produtos gênicos envolvidos no reparo do DNA danificado também são necessários para recombinação entre cromossomos homólogos, ou *crossing-over*, examinaremos agora alguns dos aspectos moleculares deste processo importante. Além disso, a recombinação geralmente, e talvez sempre, envolve alguma síntese de reparo do DNA. Assim, grande parte da informação discutida nas seções anteriores é relevante para o processo de recombinação.

Em eucariontes, o *crossing-over* está associado à formação do complexo sinaptonêmico, que se forma durante a prófase da primeira divisão meiótica. Esta estrutura é composta primariamente de proteínas e RNA. Por motivos desconhecidos, o *crossing-over* ocorre apenas raramente em machos de *Drosophila*. (O *crossing-over* ocorre em ambos os sexos da maioria das espécies. A quase ausência de *crossing-over* no sexo heterogamético é única para as *Drosophila* e algumas outras espécies.) Aqui é de interesse o fato de que o complexo sinaptonêmico está presente durante a primeira divisão meiótica da espermatogênese em *Drosophila*. Além disso, mutações no gene *c3G* de *Drosophila* eliminam tanto o *crossing-over* quanto a formação do complexo sinaptonêmico em fêmeas. Assim, o *crossing-over* e a formação do complexo sinaptonêmico parecem estar ligados. Uma pequena quantidade de síntese de DNA ocorre durante a formação do complexo sinaptonêmico, e esta síntese de DNA provavelmente está envolvida em sinapse e *crossing-over*.

Recombinação: Clivagem e Reunião de Moléculas de DNA

No Cap. 7, discutimos o experimento de Creighton e McClintock mostrando que o *crossing-over* ocorre por quebra de cromossomos parentais e reunião das partes em novas combinações. Evidência demonstrando que a recombinação ocorre por quebra e reunião também foi obtida por autoradiografia e outras técnicas. De fato, as principais características do processo de recombinação hoje estão bem estabelecidas, muito embora detalhes específicos ainda não tenham sido elucidados.

Muito do que sabemos sobre os detalhes moleculares do *crossing-over* baseia-se no estudo de **mutantes deficientes de recombinação** de *E. coli* e *S. cerevisiae*. Estudos bioquímicos destes mutantes mostraram que eles são deficientes em várias enzimas e outras proteínas necessárias para recombinação. Juntos, os resultados de estudos genéticos e bioquímicos forneceram um quadro razoavelmente completo de recombinação em nível molecular.

Muitos dos modelos atualmente populares de *crossing-over* foram derivados de um modelo proposto por Robin Holliday em 1964. O modelo de Holliday foi um dos primeiros que explicaram a maioria dos dados genéticos disponíveis na época por um mecanismo que envolve quebra, reunião e reparo das moléculas de DNA. Uma versão atualizada do modelo de Holliday é mostrada na Fig. 13.28. Este mecanismo,

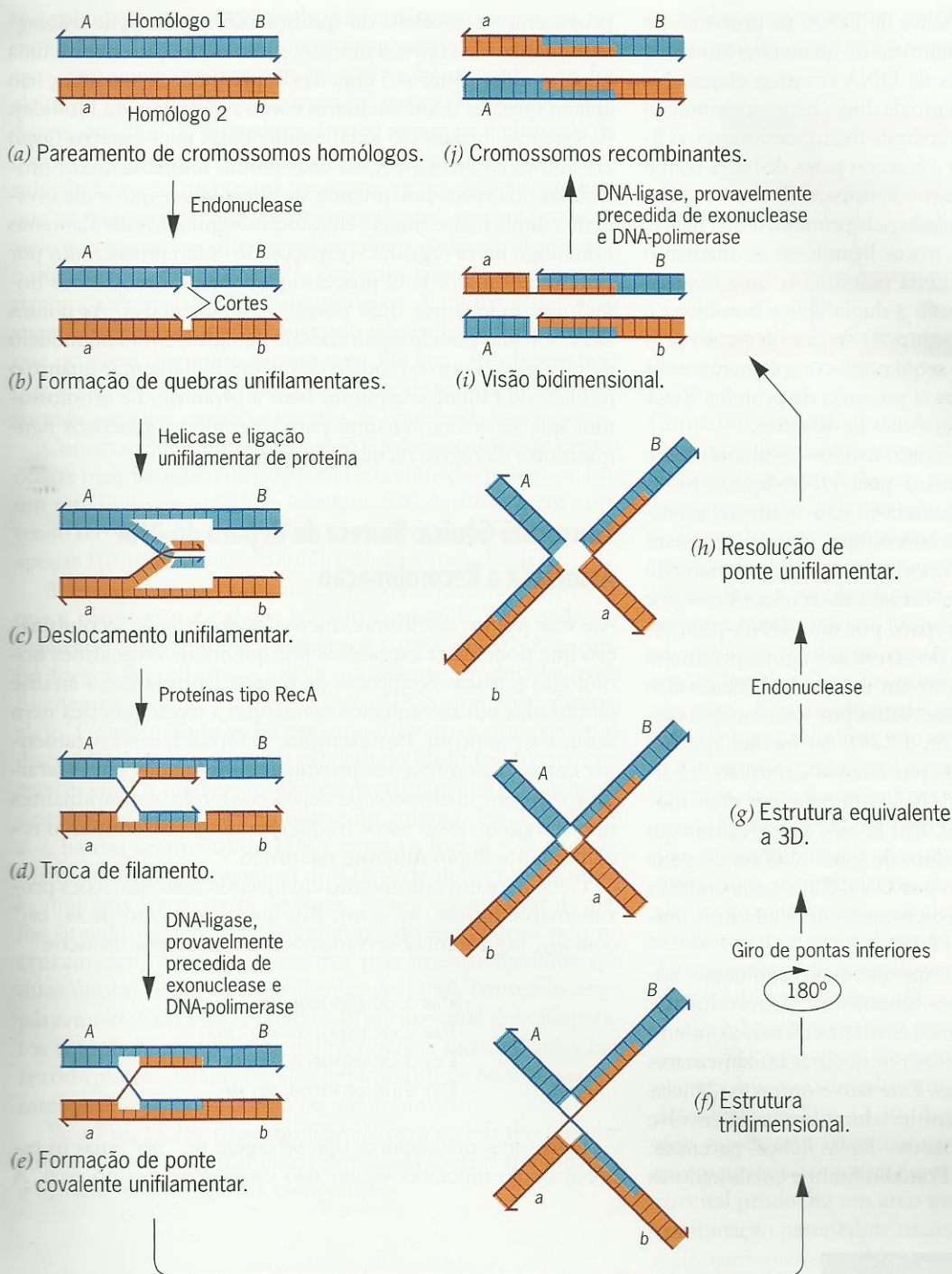


Fig. 13.28 ■ Um mecanismo para recombinação entre moléculas homólogas de DNA. A via mostrada é baseada no modelo originalmente proposto por Robin Holliday em 1964.

como muitos outros que foram invocados, começa quando uma endonuclease corta filamentos únicos de cada uma das duas moléculas de DNA parental (quebra). Segmentos dos filamentos únicos em um lado de cada corte são então deslocados de seus filamentos complementares com a ajuda de DNA-helicases e proteínas de ligação unifilamentar. As helicases deslicoidizam os dois filamentos de DNA na região adjacente a cortes unifilamentares. Em *E. coli*, o **complexo RecBCD** contém tanto atividade de endonuclease que faz quebras unifilamentares no DNA quanto atividade de DNA-

helicase que deslicoidiza os filamentos complementares de DNA na região adjacente de cada corte.

Os filamentos únicos deslocados então mudam de parceiros de pareamento, fazendo par de bases com os filamentos complementares intactos de cromossomos homólogos. Este processo é estimulado por proteínas como a **proteína RecA** de *E. coli*. Proteínas tipo RecA foram caracterizadas em muitas espécies, tanto procarióticas quanto eucarióticas. A proteína RecA e seus homólogos estimulam a **assimilação unifilamentar**, processo pelo qual um único filamento de DNA desloca

seu homólogo em uma dupla hélice de DNA. As proteínas do tipo RecA promovem trocas recíprocas de filamentos únicos de DNA entre duas duplas hélices de DNA em duas etapas. Na primeira etapa, um único filamento da dupla hélice é assimilado por um segundo, o homólogo da dupla hélice, deslocando o filamento homólogo ou idêntico e fazendo pares de bases com o filamento complementar. Na segunda etapa, o filamento único deslocado é similarmente assimilado pela primeira dupla hélice. A proteína RecA medeia essas trocas ligando-se ao filamento de DNA não pareado, ajudando na pesquisa de uma sequência de DNA homóloga e, quando a dupla hélice homóloga é encontrada, promovendo a substituição de um filamento pelo filamento não pareado. Se as sequências complementares já existem como filamentos únicos, a presença da proteína RecA aumenta a taxa de renaturação em mais de 50 vezes.

Os filamentos clivados são então unidos covalentemente em novas combinações (reunião) pela DNA-ligase. Se as quebras originais nos dois filamentos não ocorrem exatamente no mesmo sítio nos dois homólogos, algum ajuste será necessário antes que a DNA-ligase possa catalisar a etapa de reunião. Este ajuste envolve a excisão de nucleotídeos por uma exonuclease e síntese de reparo por uma DNA-polimerase. A sequência de eventos descritos até agora produzirá intermediários de recombinação em forma de X chamados **formas qui** (*chi*), que foram observadas por microscopia eletrônica em várias espécies (Fig. 13.29). As formas qui são resolvidas por quebra catalisada por enzima e reunião dos filamentos complementares de DNA para produzir duas moléculas de DNA recombinante. Em *E. coli*, as estruturas qui podem ser resolvidas pelo produto do gene *recG* ou do gene *ruvC* (reparo de dano induzido por UV). Cada gene codifica uma endonuclease que catalisa a clivagem de filamentos únicos nas junções qui (veja Fig. 13.28).

Evidências substanciais indicam que a recombinação homóloga ocorre por mais de um mecanismo, provavelmente por vários mecanismos diferentes. Em *S. cerevisiae*, as pontas das moléculas de DNA produzidas por quebras bifilamentares são altamente recombinogênicas. Este fato e outras evidências sugerem que a recombinação em leveduras em geral envolve uma quebra bifilamentar em uma das duplas hélices parentais. Assim, em 1983, Jack Szostak, Franklin Stahl e colaboradores

propuseram o **modelo da quebra bifilamentar** de *crossing-over*. De acordo com seu modelo, a recombinação envolve uma quebra bifilamentar em uma das duplas hélices parentais, não apenas quebras unifilamentares como no modelo de Holliday. As quebras iniciais são então aumentadas para espaços (*gaps*) em ambos os filamentos. As duas pontas unifilamentares produzidas no espaço bifilamentar da dupla hélice quebrada invadem a dupla hélice intacta e deslocam segmentos do filamento homólogo nesta região. Os espaços são então preenchidos por síntese de reparo. Este processo gera dois cromossomos homólogos unidos por duas pontes unifilamentares. As pontes são resolvidas por clivagem endonucleolítica, como no modelo de Holliday. Tanto o modelo de quebra bifilamentar quanto o modelo de Holliday explicam bem a produção de cromossomos que são recombinantes para marcadores genéticos flancuadores da região na qual ocorre o *crossing*.

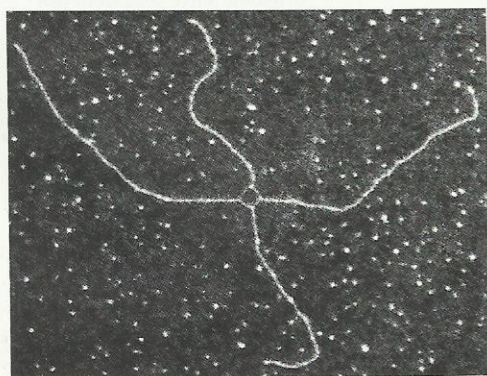
Conversão Gênica: Síntese de Reparo do DNA Associada a Recombinação

Até este ponto, discutimos apenas os eventos de recombinação que podem ser explicados por quebra de cromátides homólogas e trocas recíprocas de partes. Entretanto, a análise de tétrades em ascomicetos revela que a troca genética nem sempre é recíproca. Por exemplo, se forem feitos cruzamentos entre duas mutações proximamente ligadas em *Neurospora* e forem analisados os ascos contendo recombinantes tipo selvagem, esses ascos freqüentemente não contêm o recombinante duplo mutante recíproco.

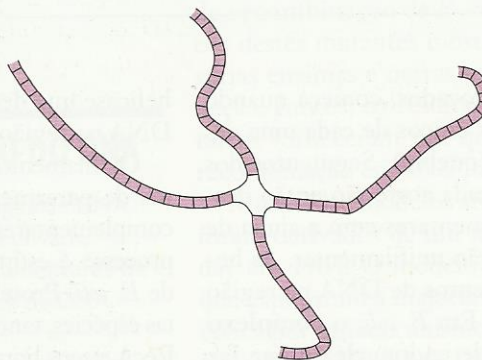
Considere um cruzamento envolvendo duas mutações proximamente ligadas, m_1 e m_2 . Em um cruzamento de $m_1 m_2^+$ com $m_1^+ m_2$, foram observados os seguintes tipos de ascos:

- Par 1 de esporos: $m_1^+ m_2$
- Par 2 de esporos: $m_1^+ m_2^+$
- Par 3 de esporos: $m_1 m_2^+$
- Par 4 de esporos: $m_1 m_2$

Estão presentes esporos tipo selvagem $m_1^+ m_2^+$, mas os esporos duplo-mutantes $m_1 m_2$ não estão presentes no asco. A



(a)



(b)

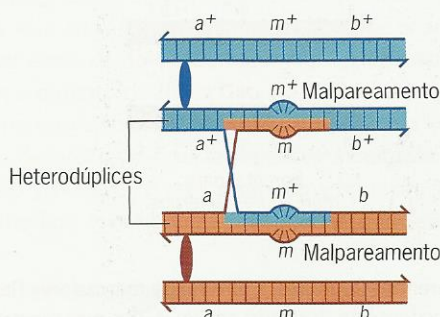
Fig. 13.29 ■ Micrografia eletrônica (a) e diagrama (b) de um intermediário de recombinação em forma de X ou estrutura qui.

recombinação recíproca produziria um cromossomo $m_1 m_2$ sempre que fosse produzido um cromossomo $m_1^+ m_2^+$. Neste asco, a proporção $m_2^+ : m_2$ é de 3:1 e não a 2:2 esperada. Um dos alelos m_2 parece ter sido “convertido” para a forma alélica m_2^+ . Assim, este tipo de recombinação não recíproca é chamada de **conversão gênica**. Podemos supor que a conversão gênica resulta de mutação, exceto que ela ocorre em frequência maior do que os eventos de mutação correspondentes, sempre produz o alelo presente no cromossomo homólogo, e não um alelo novo, e está correlacionada em 50% das vezes com a recombinação recíproca de marcadores flanqueadores. A última observação sugere fortemente que a conversão gênica resulta de eventos que ocorrem durante o *crossing-over*. De fato, acredita-se hoje que a conversão gênica resulta da síntese de reparo do DNA associada a quebra, excisão e eventos de reunião de *crossing-over*.

Com marcadores proximamente ligados, a conversão gênica ocorre mais frequentemente que a recombinação recíproca. Em um estudo do gene *his1* de levedura, 980 de 1.081 ascos contendo recombinantes *his*⁺ exibiram conversão gênica, enquanto apenas 101 mostraram recombinação recíproca clássica.

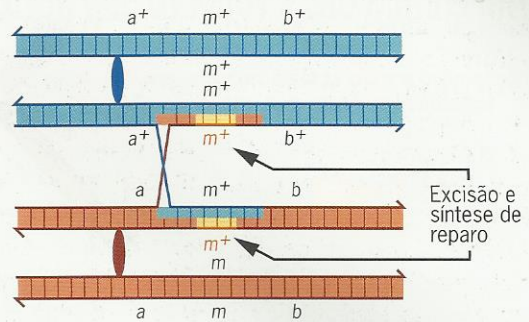
A característica mais marcante da conversão gênica é que a proporção de alelos 1:1 não é mantida. Isto pode ser explicado facilmente se segmentos curtos do DNA parental forem degradados e então novamente produzidos com filamentos moldes fornecidos pelo DNA que leva o outro alelo. Considerando os mecanismos de reparo por excisão já discutidos neste capítulo, o modelo de Holliday do *crossing-over* explica a conversão gênica para marcadores genéticos situados na vizinhança imediata do *crossing*. Na Fig. 13.28 *d-i*, há um segmento de DNA entre os loci *A* e *B* onde os filamentos complementares do DNA de dois cromossomos homólogos fazem pares de bases. Se um terceiro par de alelos situado dentro deste segmento estivesse segregando no cruzamento, estariam presentes pareamentos errados nas duas duplas hélices. As moléculas de DNA contendo esses pareamentos errados, ou alelos diferentes nos dois filamentos complementares de uma dupla hélice, são chamadas **heterodúplexes**. Tais moléculas heterodúplexes ocorrem como intermediários no processo de recombinação.

Se a Fig. 13.28e fosse modificada para incluir um terceiro par de alelos e as outras duas cromátides fossem adicionadas, a tétrade teria a seguinte composição:



Se os pareamentos errados são resolvidos mediante reparo por excisão (veja Fig. 13.27), em que os filamentos *m* são

excisados e refeitos com os filamentos complementares *m*⁺ como moldes, resultará a seguinte tétrade:



Como um resultado de replicação semiconservativa do DNA durante a divisão mitótica subsequente, esta tétrade produzirá um asco contendo seis ascospores *m*⁺ e dois ascospores *m*, a proporção de conversão gênica de 3:1.

Suponha que apenas um dos dois pareamentos errados na tétrade descrita seja reparado antes da divisão mitótica. Neste caso, a replicação semiconservativa do heterodúplex restante produzirá um homodúplex *m*⁺ e um homodúplex *m*, e o asco resultante conterá uma proporção de ascospores de 5*m*⁺:3*m*. Tais proporções de conversão gênica de 5:3 ocorrem. Elas resultam de segregação pós-meiótica (mitótica) de heterodúplexes não reparados.

Conversão gênica está associada à recombinação recíproca de marcadores flanqueadores em aproximadamente 50% das vezes. Esta correlação é bem explicada pelo modelo de Holliday de recombinação apresentado na Fig. 13.28. Se as duas cromátides recombinantes da tétrade diagramada são desenhadas de uma forma equivalente à mostrada na Fig. 13.28g, a associação da conversão gênica com a recombinação recíproca de marcadores flanqueadores pode facilmente ser explicada (Fig. 13.30). A ponte unifilamentar que conecta as duas cromátides deve ser resolvida por clivagem endonucleolítica para completar o processo de recombinação. Essa clivagem pode ocorrer seja horizontalmente ou verticalmente na forma qui desenhada na Fig. 13.30. A clivagem vertical produzirá um asco mostrando tanto conversão gênica quanto recombinação recíproca dos marcadores flanqueadores. A clivagem horizontal produzirá um asco mostrando a conversão gênica e a combinação parental de marcadores flanqueadores. Assim, se a clivagem ocorre no plano vertical em metade das vezes e no plano horizontal em metade das vezes, a conversão gênica será associada a recombinação recíproca de marcadores flanqueadores em cerca de 50% das vezes, como observado.

■ PONTOS IMPORTANTES

- *Crossing-over* envolve a quebra de moléculas homólogas de DNA e a reunião das partes em novas combinações.
- Quando marcadores genéticos estão proximamente ligados, a recombinação não recíproca, ou conversão gênica, geralmente ocorre, produzindo proporções de 3:1 de segregação de alelos.
- A conversão gênica resulta da síntese de reparo do DNA que ocorre durante o processo de recombinação.

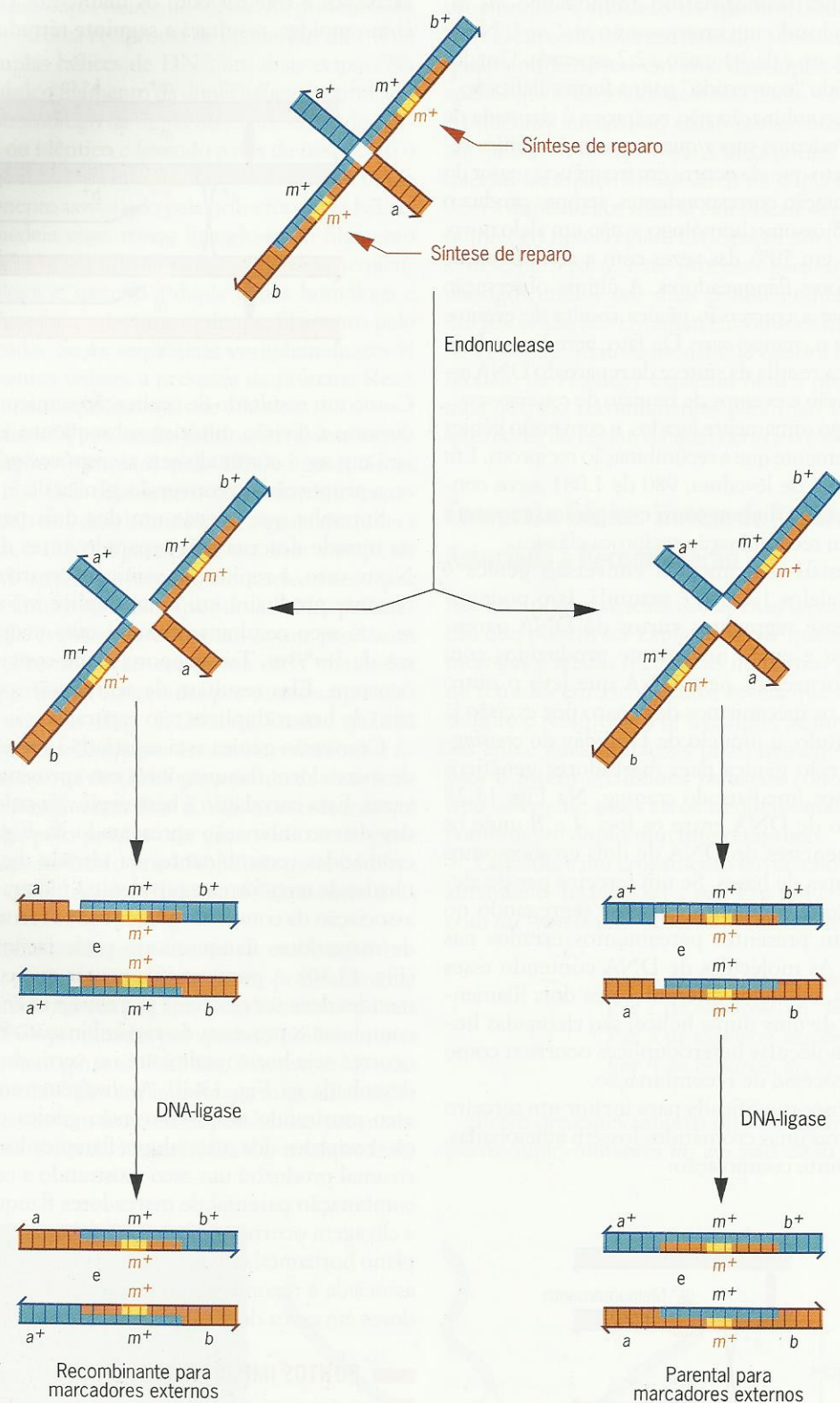


Fig. 13.30 ■ Formação de combinações recombinantes (embaixo, à esquerda) ou parentais (embaixo, à direita) de marcadores flanqueadores em associação a conversão gênica. O intermediário de recombinação no topo é equivalente ao ilustrado na Fig. 13.28g, mas mostra as cromátides com reparo de malpareamento da tétrade diagramada no texto. Esta tétrade produz um asco mostrando a conversão gênica de 3 m^+ em 1 m . A clivagem da ponte unifilar no plano vertical (à esquerda) produz o arranjo recombinante (a^+b e $a b^+$) de marcadores flanqueadores, enquanto a clivagem no plano horizontal produz o arranjo parental (a^+b^+ e $a b$) dos marcadores flanqueadores.



UM MARCO NA GENÉTICA: Muller Demonstrou que os Raios X São Mutagênicos

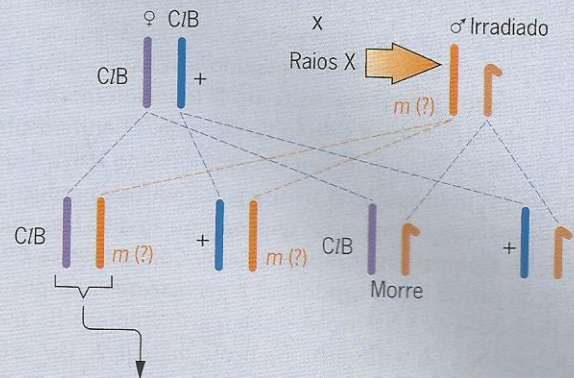
Nos anos 1880, Hugo de Vries observou o súbito aparecimento de novos fenótipos nas primulas que cresciam em seu jardim experimental. Ele chamou as novas plantas de “sports” e lhes deu nomes latinos porque pensou que fossem novas espécies. Em 1901, de Vries introduziu o termo mutação para descrever a origem destas “sports”.¹ Como ficou claro, as “sports” de de Vries resultavam de grandes mudanças na estrutura cromossômica, as translocações (veja Cap. 6).

Os novos fenótipos que resultam de mutações foram descritos e apresentados como herdáveis por muitos dos primeiros cultivadores de plantas. Entretanto, o trabalho sobre *Drosophila* no laboratório de Thomas Hunt Morgan forneceu as mais amplas evidências para os efeitos de novos alelos mutantes. As primeiras medidas quantitativas de frequência das mutações foram feitas em 1919 por Hermann J. Muller e Edgar Altenburg.² Estes pesquisadores estudaram a frequência de letais ligados ao X em *Drosophila*. A conquista de destaque nas pesquisas sobre mutação ocorreu em 1927, quando Muller demonstrou que o tratamento de espermatozoides de *Drosophila* com raios X aumentou muito a frequência de letais recessivos ligados ao X.³ (Os raios X são uma forma de radiação eletromagnética com comprimentos de onda mais curtos e maior energia que a luz visível; veja Fig. 13.20.) O estudo de Muller foi a primeira demonstração de que a mutação pode ser induzida por um fator externo. Em 1946, ele recebeu o Prêmio Nobel de Fisiologia ou Medicina por esta importante descoberta.

A demonstração inequívoca de Muller da mutagenicidade dos raios X tornou-se possível porque ele desenvolveu uma técnica simples e precisa que podia ser usada para identificar mutações letais no cromossomo X de *Drosophila*. Esta técnica, chamada de **método CIB**, é feita com fêmeas heterozigotas para um cromossomo X normal e um cromossomo X alterado, o **cromossomo CIB**, que Muller construiu especificamente para utilização em seu experimento.

O cromossomo CIB tem três componentes essenciais. (1) C, de supressor de *crossing*, refere-se a uma longa inversão que suprime a recombinação entre o cromossomo CIB e o cromossomo X estruturalmente normal em fêmeas heterozigotas. A inversão não impede *crossing* entre os dois cromossomos, mas faz com que a prole que leva cromossomos X recombinantes produzidos por *crossing-over* entre os dois cromossomos seja abortada devido a duplicações e deficiências (veja Cap. 6). (2) I refere-se a uma mutação letal recessiva no cromossomo CIB. As fêmeas homozigotas e os machos heterozigotos portadores desta mutação letal ligada ao X não são viáveis. (3) B refere-se a uma mutação que causa o fenótipo olho *bar*, uma condição na qual os grandes olhos com-

Cruzamento I: Fêmeas heterozigotas para o cromossomo CIB são cruzadas com machos irradiados.



Cruzamento II: Prole feminina CIB do cruzamento I é cruzada com machos tipo selvagem.

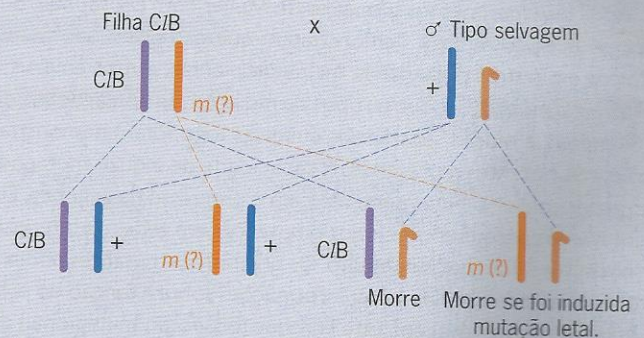


Fig. 1 ■ A técnica CIB usada por Muller para detectar mutações letais recessivas ligadas ao X em *Drosophila*. A reprodução mostrada no cruzamento II produzirá apenas uma prole feminina se um letal recessivo ligado ao X estiver presente no cromossomo X irradiado. Um terço da prole produzida pelo cruzamento II será de machos se não tiver um letal recessivo no cromossomo X irradiado. Assim, o registro de mutações letais simplesmente envolve triagem da prole do cruzamento II quanto à presença ou à ausência de machos.

postos de moscas tipo selvagem são reduzidos a olhos estreitos em forma de barra. Como B é parcialmente dominante, as fêmeas heterozigotas para o cromossomo CIB podem ser identificadas prontamente. Tanto a mutação letal recessiva (I) quanto o olho *bar* (B) estão situados dentro do segmento invertido do cromossomo CIB.

Muller irradiou machos e os cruzou com fêmeas CIB (Fig. 1). Todas as fêmeas com olho *bar* deste cruzamento levavam o cromossomo CIB do genitor feminino e o cromossomo X irradiado do genitor masculino. Como toda a população de células reprodutivas dos machos foi irradiada, cada filha com olho *bar* tinha um cromossomo X que potencialmente sofreu mutação. Estas filhas com olhos *bar* foram então individualmente cruzadas (em culturas

¹de Vries, H. 1901. *Die Mutationstheorie*, Vol. 1, Viet, Leipzig.

²Muller, H. J., e E. Altenburg. 1919. The rate of change of hereditary factors in *Drosophila*. *Proc. Soc. Exp. Biol. Med.* 17: 10-14.

³Muller, H. J. 1927. Artificial transmutation of the gene. *Science* 66: 84-87.



UM MARCO NA GENÉTICA (continuação)

separadas) com machos tipo selvagem. Se o cromossomo X irradiado levado por uma fêmea com olhos *bar* tivesse adquirido um letal ligado ao X, toda a prole do cruzamento seria de fêmeas. Os machos hemizigotos para o cromossomo *ClB* morreriam devido ao letal recessivo (*l*) que este cromossomo leva. Além disso, os machos hemizigotos para o cromossomo X irradiado morreriam se um letal recessivo tivesse sido induzido nele. O cruzamento das filhas com olhos *bar* que levam um cromossomo X irradiado no qual nenhuma mutação letal tivesse sido induzida produziria uma prole de machos e fêmeas em uma proporção de 2:1 (apenas machos com o cromossomo *ClB* morreriam). Com a técnica *ClB*, a detecção de novos letais ligados ao X recessivos induzidos é precisa e livre de erro. Ela envolve nada mais complexo do que registrar a presença ou a ausência da prole masculina. Por este procedimento, Muller foi capaz de demonstrar um aumento de 150 vezes na frequência de letais ligados ao X após tratar moscas fêmeas com raios X.

Após os experimentos pioneiros de Muller com *Drosophila*, outros pesquisadores logo demonstraram que os raios X são mutagênicos em outros organismos, incluindo plantas, outros animais e micróbios. Além disso, outros tipos de radiação eletromagnética de alta energia e muitas substâncias químicas logo foram demonstradas como potentes mutágenos. A capacidade de induzir mutações em genes contribuiu imensamente para o progresso da genética. Ela permitiu aos pesquisadores induzir mutações em genes de interesse e “nocautear” seu funcionamento. Os organismos mutantes podiam então ser estudados para obter informações sobre a função do produto gênico tipo selvagem. Este enfoque (dissecção mutacional) de processos biológicos demonstrou ser um poderoso instrumento na análise de muitos processos biológicos.

QUESTÕES PARA DISCUSSÕES

1. Muller demonstrou que os raios X induzem mutações em *Drosophila*. Estudos subsequentes em muitos laboratórios mostraram que os raios X são mutagênicos em todos os organismos, incluindo seres humanos. Entretanto, os raios X também são uma importante ferramenta diagnóstica em medicina e odontologia. Quando você se submete a um procedimento de raios X diagnóstico, o profissional médico fornece um avental de chumbo para evitar a irradiação de suas células reprodutivas. Isto as protege e diminui a chance de novas mutações que possam ser transmitidas para sua prole. Mas os raios X induzem mutações em células somáticas*, e algumas dessas mutações aumentam a chance de câncer no futuro. Que regulamentações se aplicam ao uso de raios X para fins diagnósticos? Que fatores devem ser considerados ao estabelecer tais regulamentações?
2. Além dos raios X, outras formas de radiação de alta energia tais como os raios gama, os raios cósmicos e as partículas liberadas pelo decaimento de substâncias radioativas usadas em bombas nucleares e reatores nucleares são altamente mutagênicas. Em 1986, os riscos potenciais de reatores nucleares tornaram-se realidade quando a usina nuclear perto de Chernobyl na Ucrânia explodiu e contaminou a cidade vizinha. Embora os reatores nucleares sejam uma fonte econômica de eletricidade, o potencial de acidentes como o de Chernobyl permanece. Quais são os prós e os contras do uso de reatores nucleares para fornecer energia econômica e “ecologicamente segura”?

*N.T.: Embora em geral não seja feito, é necessário proteger também a tireóide.

Exercícios Básicos Ilustram a análise genética básica.

1. Considere o papel da mutação na evolução. Uma espécie poderia evoluir na ausência de mutações?

Resposta: Não. Mutação é uma primeira etapa essencial no processo evolutivo; é a principal fonte de toda a nova variação genética. Os mecanismos de recombinação produzem novas combinações dessa variação genética, e a seleção natural (ou artificial) preserva as combinações que produzem organismos mais bem adaptados aos ambientes nos quais eles vivem. Sem mutação, não poderia haver evolução.

2. Considere um segmento curto de um gene tipo selvagem com a seguinte sequência de pares de nucleotídeos:

```

5'-ATG TCC GCA TGG GGA-3'
   |||  |||  |||  |||  |||
3'-TAC AGG CGT ACC CCT-5'
  
```

A transcrição deste segmento gênico produz a seguinte sequência de nucleotídeos do mRNA:

5'-AUG UCC GCA UGG GGA-3'

e a tradução deste mRNA produz a sequência de aminoácidos:

metionina-serina-alanina-triptofano-glicina

Se ocorrer substituição de um único par de nucleotídeos neste gene, trocando G-C na posição 7 por A-T, que efeito esta mutação tem no polipeptídeo produzido por este gene?

Resposta: O mRNA produzido pelo segmento do gene com a mutação agora será:

5'-AUG UCC ACA UGG GGA-3'

e codificará a sequência de aminoácidos:

metionina-serina-treonina-triptofano-glicina

Note que o terceiro aminoácido do polipeptídeo mutante é treonina em vez de alanina como no polipeptídeo tipo selvagem. Assim, esta substituição de par de bases, como a maioria das substituições de par de bases, resulta em uma única substituição de aminoácido no polipeptídeo codificado pelo gene.

3. Se ocorre uma única substituição de par de nucleotídeos no segmento gênico mostrado no Exercício 2, mudança de G-C na posição 12 para A-T, que efeito a mutação terá no polipeptídeo produzido por este gene?

Resposta: A sequência do mRNA resultante será:

5'-AUG UCC ACA UGA GGA-3'
códon de fim

com o quarto códon mudado de UGG, um códon de triptofano, para UGA, um dos três códons de fim de cadeia. Como resultado, o polipeptídeo mutante será terminado prematuramente nesta posição, gerando uma proteína truncada.

4. Se um único par de bases A-T é inserido entre os pares de nucleotídeos 6 e 7 no segmento gênico mostrado no Exercício 2, que efeito esta mudança tem no polipeptídeo especificado por este gene?

Resposta: A sequência de nucleotídeos do mRNA especificada pelo segmento gênico mutante será:

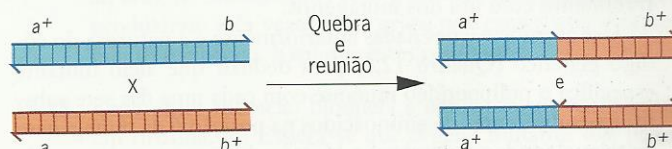
5'-AUG UCC AGC AUG GGG A-3'

e o polipeptídeo produzido pelo mRNA alterado será:

metionina-serina-serina-metionina-glicina
sequência alterada de aminoácidos

A inserção de um par de bases irá alterar a matriz de leitura do mRNA (trinucleotídeos lidos como códons) distal ao sítio da mutação. Como resultado, todos os aminoácidos especificados pelos códons posteriores ao sítio de inserção serão mudados, produzindo uma proteína anormal (em geral não funcional). Em muitos casos, uma inserção mudará o códon de término para a matriz de leitura apropriada para tradução, causando a produção de um polipeptídeo truncado.

5. Se as duas moléculas de DNA mostradas a seguir, onde a cabeça da seta indica a ponta 3' de cada filamento, sofrerem *crossing* por quebra e reunião, ambos os recombinantes mostrados serão produzidos com frequências iguais?



Resposta: Não. Durante a recombinação, apenas os filamentos de DNA com a mesma polaridade podem ser unidos. O segundo recombinante não será produzido.

Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. Charles Yanofsky isolou um grande número de mutantes auxotróficos de *E. coli* que só podiam crescer em meio contendo o aminoácido triptofano. Como tais mutantes podem ser identificados? Se um auxotrófico específico de triptofano resultar de uma mutação induzida por ácido nitroso, ele pode ser induzido a reverter para prototrófico por tratamento com 5-bromouracil (5-BU)?

Resposta: A cultura de bactérias mutagenizadas deve ser em um meio contendo triptofano de modo que os mutantes desejados possam sobreviver e se reproduzir. As bactérias devem então ser diluídas, plaqueadas em meio de ágar contendo triptofano e incubadas até que sejam produzidas colônias visíveis. As colônias em seguida são transferidas para placas sem triptofano pela técnica de plaqueamento réplica desenvolvida pelos Lederberg (veja Fig. 13.2). Os auxotróficos de triptofano desejados crescerão em placas contendo triptofano mas não em placas réplica sem triptofano. Como ácido nitroso e 5-BU produzem mutações de transição em ambas as direções, A-T \leftrightarrow G-C, qualquer mutação induzida por ácido nitroso deve ser induzida a retromutar com 5-BU.

2. Suponha que você descobriu recentemente uma nova espécie de bactérias e a denominou *Escherichia mutaphilium*. Durante o

último ano, você estudou o gene *mutA* e seu produto polipeptídico, a enzima trinucleotídeo-mutagenase, nesta bactéria. Foi demonstrado que *E. mutaphilium* usa o código genético estabelecido, quase universal, e se comporta como *Escherichia coli* em todos os outros aspectos relevantes da genética molecular.

O sexto aminoácido da ponta amino do trinucleotídeo tipo selvagem mutagenase é histidina, e o gene tipo selvagem *mutA* tem a sequência de trincas de pares de nucleotídeos

3'-G T A-5'
... ..
5'-C A T-3'

na posição correspondente ao sexto aminoácido do produto gênico. Também foram caracterizados sete mutantes isolados independentemente com substituições de um único par de nucleotídeos dentro desta trinca. Além disso, as mutagenases de trinucleotídeos mutantes foram todas purificadas e sequenciadas. Todas as sete são diferentes: elas contêm, respectivamente, glutamina, tirosina, asparagina, ácido aspártico, arginina, prolina e leucina como o sexto aminoácido a partir da ponta amino.

Os mutantes *mutA1*, *mutA2* e *mutA3* não se recombinam uns com os outros, mas cada um irá recombinar-se com cada um dos outros quatro mutantes (*mutA4*, *mutA5*, *mutA6* e *mutA7*) para produzir recombinantes verdadeiramente tipo selvagem. Similarmente, os mutantes *A4*, *A5* e *A6* não se recombinam uns com os outros mas produzirão verdadeiros recombinantes tipo selvagem em cruzamentos com cada um dos outros quatro mutantes. Finalmente, cruzamentos entre *mutA1* e *mutA7* produzem cerca de até o dobro de recombinantes do tipo selvagem, como cruzamentos entre *mutA6* e *mutA7*.

Os mutantes *A1* e *A6* são induzidos a retromutar para o tipo selvagem por tratamento com 5-bromouracil (5-BU), enquanto os mutantes *A2*, *A3*, *A4*, *A5* e *A7* não são induzidos a retromutar por tratamento com 5-BU. Os mutantes *A2* e *A4* crescem lentamente em meio mínimo, enquanto os mutantes *A3* e *A5* levam mutações nulas (produzindo produtos gênicos completamente inativos) e são incapazes de crescer em meio mínimo. Esta diferença foi usada para selecionar eventos mutacionais dos genótipos *mutA3* e *mutA5* para os genótipos *mutA2* e *mutA4*. Os mutantes *A3* e *A5* podem ser induzidos a mutar para *A2* e *A4* respectivamente, por tratamento com 5-bromouracil ou hidroxilamina. Entretanto, o mutante *A3* não pode ser induzido a mutar para *A4* nem *A5* para *A2*, por tratamento com um dos mutágenos.

Use as informações dadas anteriormente e a natureza do código genético (Quadro 12.1) para deduzir que alelo mutante específica o polipeptídeo mutante com cada uma das sete substituições diferentes de aminoácidos na posição 6 da mutagenase de trinucleotídeo e descreva a lógica subjacente a cada uma de suas deduções.

Resposta: As seguintes deduções podem ser tiradas das informações dadas:

- (1) O códon tipo selvagem para His deve ser CAU com base na sequência de pares de nucleotídeos do gene.
- (2) Os códons dos sete aminoácidos encontrados na posição 6 nos polipeptídeos mutantes devem estar conectados a CAU por uma única mudança de bases porque os mutantes são todos derivados do tipo selvagem por uma única substituição de par de nucleotídeos. Assim, a redundância do código genético não é um fator para deduzir atribuições de códons específicos.
- (3) Devido à natureza do código genético, especificamente a redundância da terceira posição (3') em cada códon, existem três substituições possíveis de aminoácidos devidas a substituições de uma base (causadas por substituições de um único par de bases no DNA) em cada uma das primeiras duas posições (a base 5' e a base média), mas apenas uma mudança possível de aminoácidos devida a uma mudança de uma base na posição 3 (a base em 3' no códon). Para facilitar a discussão, as posições nos três pares de nucleotídeos na trinca em consideração serão referidas como posição 1 (correspondendo à base 5' no códon), posição 2 (o par médio de nucleotídeos) e posição 3 (correspondendo à base 3' no códon do mRNA).
- (4) Como *A1*, *A2* e *A3* não se recombinam entre si, devem resultar de substituições de pares de bases na mesma posição na trinca, ou na posição 1 ou na posição 2. O mesmo é verdade para *A4*, *A5* e *A6*. Como *A7* se recombina com cada um dos outros seis alelos mutantes, ele deve resultar da substituição de um único par de bases na posição 3 que leva a uma mudança de aminoácidos.
- (5) O único aminoácido com códons conectados ao códon de His CAU por mudanças de uma base na posição 3 é Gln (códon

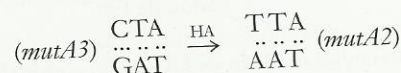
CAA e CAG). Assim, o polipeptídeo *mutA7* deve ter glutamina como o sexto aminoácido.

(6) Como *mutA7* (a substituição na terceira posição) produz cerca do dobro das recombinações tipo selvagem em cruzamentos com *mutA1* que nos cruzamentos com *mutA6*, a substituição *A1* deve ser na posição 1 e a substituição *mutA6* deve ser na posição 2. Combinada com (4) acima, isto coloca as substituições *A2* e *A3* na posição 1 e as substituições *A4* e *A5* na posição 2.

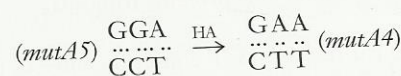
(7) Como *mutA1* e *mutA6* são induzidas a reverter ao tipo selvagem por 5-BU, elas devem estar conectadas às trincas de pares de nucleotídeos que codificam His por mutações de transição, isto é,



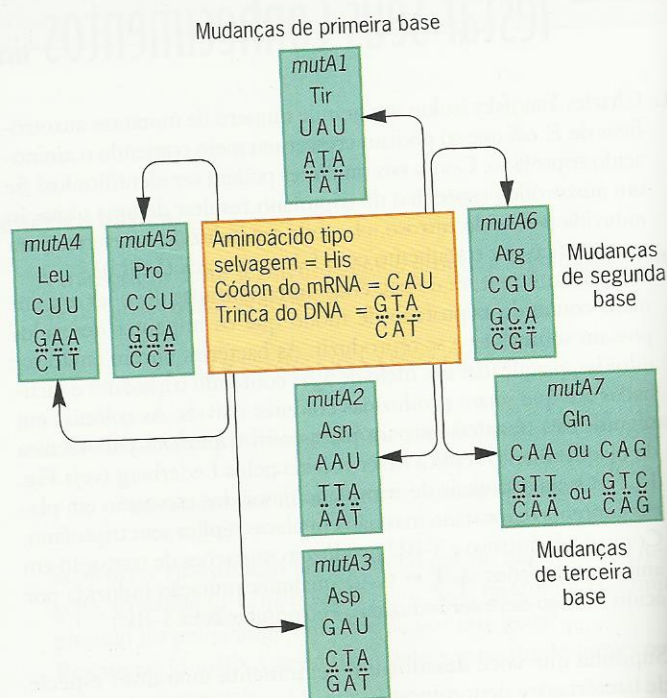
(8) Como *mutA3* e *mutA5* são induzidas a mutar para *mutA2* e *mutA4*, respectivamente, pela hidroxilamina, *A3* deve estar conectada a *A2* e *A5* a *A4*, especificamente por transições G-C → A-T, isto é,



e



Coletivamente, estas deduções estabelecem que as correlações seguintes entre aminoácidos, códons e trincas de pares de nucleotídeos estão presentes na posição de interesse nos polipeptídeos da mutagenase de trinucleotídeos, mRNA, e nos genes nos sete mutantes diferentes:



Questões e Problemas

Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 13.1. Identifique as seguintes mutações de ponto representadas no DNA e no RNA como (1) transições, (2) transversões ou (3) mudança de matriz de leitura. (a) A para G; (b) C para T; (c) C para G; (d) T para A; (e) UAU ACC UAU para UAU AAC CUA; (f) UUG CUA AUA para UUG CUG AUA.
- 13.2. De todas as possíveis mutações de sentido trocado que podem ocorrer em um segmento de DNA que codifica o aminoácido triptofano, qual a proporção de transversões para transições se ocorrerem todas as possíveis substituições de pares de bases na mesma frequência?
- 13.3. Tanto mutações letais quanto visíveis devem ocorrer em moscas-das-frutas que são submetidas a irradiação. Cite um método para detectar (a) letais ligados ao X e (b) mutações visíveis ligadas ao X em drosófilas irradiadas.
- 13.4. Como podem ser detectadas mutações em bactérias que causam resistência a determinada droga? Como pode ser determinado se uma certa droga causa mutações ou se apenas ela identifica mutações já presentes no organismo pesquisado?
- 13.5. As taxas de mutações espontâneas publicadas para humanos são geralmente mais altas que aquelas para bactérias. Isto indica que os genes individuais de humanos mutam mais frequentemente que os de bactérias? Explique.
- 13.6. Uma condição pré-cancerosa (polipose intestinal) em uma determinada família humana é causada por um único gene dominante. Entre os descendentes de uma mulher que morreu com câncer de cólon, 10 pessoas morreram com o mesmo tipo de câncer e 6 agora têm polipose intestinal. Todos os outros ramos do grande heredograma foram cuidadosamente examinados, e nenhum caso foi encontrado. Sugira uma explicação para a origem do gene defeituoso.
- 13.7. A distrofia muscular juvenil em humanos depende de um gene recessivo ligado ao X. Em um estudo intensivo, 33 casos foram encontrados em uma população de cerca de 800.000 pessoas. Os pesquisadores estavam convictos de que tinham encontrado todos os casos que estavam bem avançados para serem detectados na época em que o estudo foi feito. Os sintomas da doença foram manifestados apenas em homens. A maioria daqueles com a doença morreram em idade jovem, e nenhum viveu além dos 21 anos de idade. Geralmente, apenas um caso foi detectado em uma família, mas às vezes dois ou três casos ocorreram na mesma família. Sugira uma explicação para a ocorrência esporádica da doença e a tendência do gene de persistir na população.
- 13.8. Os produtos que resultam de mutações somáticas, tais como a laranja de umbigo e a maçã Delicious, tornaram-se generalizados na citricultura e em plantios de maçãs. Entretanto, as características que resultam de mutações somáticas raramente são mantidas em animais. Por quê?
- 13.9. Se uma única ovelha de pernas curtas ocorrer em um rebanho, sugira experimentos para determinar se as pernas curtas são o resultado de uma mutação ou um efeito ambiental. Se decorrentes de uma mutação, como se pode determinar se a mutação é dominante ou recessiva?
- 13.10. Como podem enzimas tais como a DNA-polimerase estar envolvidas no modo de ação tanto de genes mutadores quanto de antimutadores (genes mutantes que aumentam e diminuem, respectivamente, as taxas de mutação)?
- 13.11. Como as taxas de mutações espontâneas podem ser otimizadas pela seleção natural?
- 13.12. Um gene mutador *Dt* em milho aumenta a taxa na qual o gene para aleurona incolor (*a*) muta para o alelo dominante (*A*), que produz aleurona colorida. Quando são feitos cruzamentos recíprocos (isto é, genitores de sementes *dt/dt*, *a/a* × *Dt/Dt*, *a/a* e genitores de sementes *Dt/Dt*, *a/a* × *dt/dt*, *a/a*), os cruzamentos com genitores de sementes *Dt/Dt* produziram três vezes mais grãos por espiga que o cruzamento recíproco. Explique estes resultados.
- 13.13. Uma única mutação bloqueia a conversão de fenilalanina em tirosina. (a) Espera-se que o gene mutante seja pleiotrópico? (b) Explique.
- 13.14. Como podem ser diferenciadas a hemoglobina normal (hemoglobina A) e a hemoglobina S?
- 13.15. Se CTT é uma trinca de DNA (filamento transcrito de DNA) que especifica ácido glutâmico, que alterações de trinca de bases do DNA e do mRNA codificariam valina e lisina na posição 6 da cadeia de β -globina?
- 13.16. Por que a anemia falciforme é denominada uma doença molecular?
- 13.17. Supondo que a cadeia de β -globina e a cadeia de α -globina têm um ancestral comum, que mecanismos podem explicar as diferenças que hoje existem entre estas duas cadeias? Que mudanças nos códons do DNA e do mRNA respondem pelas diferenças que resultaram em aminoácidos diferentes em posições correspondentes?
- 13.18. Em determinada linhagem de bactérias, todas as células geralmente são mortas quando está presente uma concentração específica de estreptomicina no meio. Ocorrem mutações que conferem resistência a estreptomicina. Os mutantes resistentes a estreptomicina são de dois tipos: alguns podem viver com ou sem estreptomicina; outros não podem sobreviver a menos que esta droga esteja presente no meio. Considerando uma linhagem sensível a estreptomicina desta espécie, imagine um procedimento experimental pelo qual possam ser estabelecidas linhagens resistentes a estreptomicina dos dois tipos.
- 13.19. Uma coleção de moscas-das-frutas foi tratada com 1.000 roentgens (r) de raios X. O tratamento com raios X aumen-

tou a taxa de mutação de um determinado gene em 2%. Que aumento de porcentagem na taxa de mutação deste gene seria esperado se tal coleção de moscas fosse tratada com doses de raios X de 1.500 r, 2.000 r e 3.000 r?

13.20. Por que a frequência de quebras cromossômicas induzidas por raios X varia de acordo com a dose total e não com a taxa em que são induzidas?

13.21. Um reator superaquece e produz trítio radioativo (H^3), iodo radioativo (I^{131}) e xenônio radioativo (Xn^{133}). Por que devemos ficar mais preocupados com o iodo radioativo do que com os outros dois isótopos radioativos?

13.22. Uma pessoa envolveu-se em um acidente e recebeu 50 roentgens (r) de raios X de uma vez. Outra pessoa recebeu 5 r em cada um dos 20 tratamentos. Supondo que não exista efeito de intensidade, que número proporcional de mutações seria esperado em cada pessoa?

13.23. Como a luz ultravioleta produz mutações reversíveis?

13.24. Como o ácido nitroso induz mutações? Que resultados finais específicos devem ser esperados no DNA e no mRNA em decorrência do tratamento de vírus com o ácido nitroso?

13.25. Alterações mutacionais induzidas pelo ácido nitroso mais provavelmente são transições ou transversões?

13.26. Você está triando três pesticidas novos quanto a potencial mutagenicidade usando o teste de Ames. Duas linhagens *his⁻* resultantes de mudança de matriz de leitura ou de uma mutação de transição foram usadas para produzir os seguintes resultados (números de colônias revertentes).

Linhagem 1

	Controle de Mutante de Transição (sem substâncias químicas)	Mutante de Transição + Substâncias Químicas	Mutante de Transição + Substâncias Químicas + Enzimas de Fígado de Rato
Pesticida nº 1	21	180	19
Pesticida nº 2	18	19	17
Pesticida nº 3	25	265	270

Linhagem 2

	Controle de Mutante por Matriz de Matriz de Leitura (sem substâncias químicas)	Mutante de Matriz de Leitura + Substâncias Químicas	Mutante de Matriz de Leitura + Substâncias Químicas + Enzimas de Fígado de Rato
Pesticida nº 1	5	4	5
Pesticida nº 2	7	5	93
Pesticida nº 3	6	9	7

Que tipos de mutações, se alguma, estes três pesticidas induzem?

13.27. Como a ação e o efeito mutagênico do 5-bromouracil diferem daqueles do ácido nitroso?

13.28. Sydney Brenner e A. O. W. Stretton observaram que mutações sem sentido não terminam a síntese de polipeptídeos no gene *rII* do bacteriófago T4 quando tais mutações foram localizadas dentro de um intervalo da sequência de DNA no qual uma única inserção de nucleotídeo foi feita em uma ponta e uma única deleção de nucleotídeo foi feita na outra. Como este achado pode ser explicado?

13.29. Seymour Benzer e Ernst Freese compararam mutantes espontâneos e induzidos por 5-bromouracil no gene *rII* do bacteriófago T4. O mutágeno aumentou a taxa de mutação (*rII⁺* → *rII*) várias centenas de vezes acima da taxa de mutação espontânea. Quase todos os mutantes (98%) induzidos por 5-bromouracil podiam ser induzidos a reverter para o tipo selvagem (*rII* → *rII⁺*) por tratamento com 5-bromouracil, mas apenas 14% dos mutantes espontâneos podiam ser induzidos a reverter para o tipo selvagem com este tratamento. Discuta o motivo deste resultado.

13.30. Como as alterações induzidas por acridina no DNA resultam em proteínas inativas?

Use as atribuições conhecidas de códon-aminoácido fornecidas no Cap. 12 para resolver os problemas seguintes.

13.31. Mutações nos genes que codificam as subunidades α e β de hemoglobina provocam doenças sangüíneas tais como talassemias e anemia falciforme. Você encontrou uma família na China na qual alguns membros sofrem de uma nova forma genética de anemia. As sequências de DNA da ponta 5' do filamento não-molde do DNA normal e mutante codificantes da subunidade α de hemoglobina são as seguintes:

Normal 5'-ACGTTATGCCGTACTGCCAGCTAACT-GCTAAAGAACAATTA.....-3'
Mutante 5'-ACGTTATGCCCGTACTGCCAGCTAACT-GCTAAAGAACAATTA.....-3'

- Que tipo de mutação está presente no gene mutante de hemoglobina?
- Quais são os códons na parte traduzida do mRNA transcrito dos genes normal e mutante?
- Quais são as sequências de aminoácidos dos polipeptídeos normal e mutante?

13.32. O bacteriófago MS2 leva sua informação genética no RNA. Seu cromossomo é análogo a uma molécula poligênica de mRNA em organismos que estocam sua informação genética no DNA. O minicromossomo MS2 codifica 4 polipeptídeos (isto é, ele tem quatro genes). Um desses quatro genes codifica a capa de proteína do MS2, um polipeptídeo com 129 aminoácidos de tamanho. Toda a sequência de nucleotídeos no RNA do MS2 é conhecida. O códon 112 do gene da proteína da capa é CUA, que especifica o aminoácido leucina. Se você fosse tratar uma população replicante do bacteriófago MS2 com o mutágeno 5-bromouracil, que

substituições de aminoácidos você esperaria ser induzida na posição 112 da capa de proteína de MS2 (isto é, Leu → outro aminoácido)? (Nota: o RNA do bacteriófago MS2 replica-se usando um filamento complementar de RNA e pareamento de bases como DNA.)

- 13.33. Seria esperado que as diferentes substituições de aminoácidos induzidas por 5-bromouracil na posição 112 do polipeptídeo da capa que você indicou no Problema 13.32 ocorressem com iguais frequências? Caso sim, por quê? Caso não, por que não? Qual, se alguma, ocorre mais frequentemente?
- 13.34. Tais mutações ocorreriam se uma suspensão não replicante do fago MS2 fosse tratada com 5-bromouracil?
- 13.35. Lembre que o ácido nitroso desamina adenina, citosina e guanina (adenina → hipoxantina, que faz par de bases com citosina; citosina → uracil, que faz par de bases com adenina; e guanina → xantina, que faz par de bases com citosina). Você esperaria que o ácido nitroso induzisse qualquer mutação que resultasse na substituição de outro aminoácido por uma glicina no polipeptídeo tipo selvagem (isto é, glicina → outro aminoácido) se a mutagênese fosse feita em uma suspensão de bacteriófago T4 maduro (não replicante)? (Nota: Após o tratamento mutagênico da suspensão de fago, o ácido nitroso é removido. Permite-se então que os fagos tratados infectem *E. coli* para expressar quaisquer mutações induzidas.) Assim sendo, por qual mecanismo? Caso não, por que não?
- 13.36. Tendo em mente a natureza conhecida do código genético, a informação dada sobre o fago MS2 no Problema 13.32 e as informações que você aprendeu sobre o ácido nitroso no Problema 13.35, você esperaria que o ácido nitroso induzisse quaisquer mutações que resultassem em substituições de aminoácidos do tipo glicina → outro aminoácido se a mutagênese fosse feita em uma suspensão de bacteriófagos MS2 maduros (não replicantes)? Caso sim, por que mecanismo? Caso não, por que não?
- 13.37. Você esperaria que o ácido nitroso induzisse uma frequência mais alta de substituições Tir → Ser ou Tir → Cis? Por quê?

- 13.38. Quais das seguintes substituições de aminoácidos você deveria esperar ser induzidas por 5-bromouracil com a frequência mais alta? (a) Met → Leu; (b) Met → Tre; (c) Lis → Tre; (d) Lis → Gln; (e) Pro → Arg; ou (f) Pro → Gln? Por quê?

- 13.39. A sequência tipo selvagem de uma parte de uma proteína é

NH₂-Trp-Trp-Trp-Met-Arg-Glu-Trp-Tre-Met

Cada mutante no quadro seguinte difere do tipo selvagem por uma única mutação de ponto. Usando esta informação, determine a sequência de mRNA codificante do polipeptídeo tipo selvagem. Se houver mais de um nucleotídeo possível, cite todas as possibilidades.

Mutante Sequência de Aminoácidos do Polipeptídeo

1	Trp-Trp-Trp Met
2	Trp-Trp-Trp-Met-Arg-Asp-Trp-Tre-Met
3	Trp-Trp-Trp-Met-Arg-Lis-Trp-Tre-Met
4	Trp-Trp-Trp-Met-Arg-Glu-Trp-Met-Met

- 13.40. Sabe-se que os corantes acridina tais como a proflavina induzem primariamente adições ou deleções de um só par de bases. Suponha que a sequência de nucleotídeos tipo selvagem no mRNA produzido por um gene é

5'-AUGCCCUUUGGGAAAGGGUUCUUCCCUAA-3'

Suponha também que seja induzida uma mutação neste gene pela proflavina e, subsequentemente, seja induzido um revertente desta mutação com proflavina, resultante de uma mutação supressora em segundo sítio dentro do mesmo gene. Se a sequência de aminoácidos do polipeptídeo codificado por este gene na linhagem revertente (mutante duplo) é

NH₂-Met-Pro-Fen-Gli-Glu-Arg-Fen-Pro-COOH

qual seria a sequência de nucleotídeos mais provável no mRNA deste gene no revertente (mutante duplo)?

Definições do Gene

Tópicos

- Evolução do Conceito de Gene: Resumo
- Evolução do Conceito de Gene: Função
- Evolução do Conceito de Gene: Estrutura
- Uma Definição Genética de Gene
- O Locus *rII* do Bacteriófago T4
- Genes Dentro de Genes no Bacteriófago Φ X174
- Relações do Complexo Gene-Proteína



Uma corça de pigmentação normal com sua mãe albina.

O Que É Vida?

Em um artigo sobre pesquisa genética nos anos que antecederam a descoberta da estrutura do DNA, James Watson escreveu:

*Como estudante de graduação em Chicago, eu já havia decidido pela genética muito embora meu treinamento formal nela fosse insuficiente e a maior parte de meu curso refletisse um interesse de infância pela história natural. A genética de populações a princípio me intrigava, mas no momento em que li *What Is Life?*¹ de Schrödinger fiquei direcionado para descobrir o segredo do gene.²*

¹Schrödinger, E. 1945. *What Is Life? The Physical Aspect of the Living Cell*. Cambridge University Press, London.

²Watson, J. D. 1966. Growing up in the phage group. Em *Phage and the Origins of Molecular Biology* (J. Cairns, G. S. Stent e J. D. Watson, eds.), pp. 239-245. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Watson não foi o único cientista a ser influenciado pelo pequeno livro de Erwin Schrödinger. *What Is Life?* foi escrito por este físico austríaco, famoso por seu trabalho em mecânica quântica, para explorar questões biológicas fundamentais. Entre outras coisas, ele inclui uma discussão sobre como a irradiação da matéria viva pode ser usada para determinar o tamanho de um gene. Schrödinger sabia da demonstração de Muller de que os raios X induzem mutações e ficou curioso quanto às idéias de seu colega físico Max Delbrück sobre o mecanismo subjacente. Delbrück colaborou com o geneticista russo Nikolaj V. Timofeeff-Ressovsky para analisar os efeitos da radiação e para desenvolver um modelo de “mecânica quântica” de um gene. A discussão de seu trabalho no *What Is Life?* foi lida por outros físicos, e alguns deles, com a curiosidade despertada, decidiram abandonar a física e seguir a carreira em genética. Um dos que mudaram foi um jovem americano chamado Seymour Benzer. Em um tributo a Max Delbrück (e, indiretamente, a Erwin Schrödinger), Benzer escreveu:

Delbrück entrou primeiro em minha vida na forma do título de um capítulo “Delbrück’s Model” no livro de Schrödinger What Is Life? Li este livro em uma idade impressionável, quando ainda era um estudante de graduação de física do estado sólido pré-transistor na Purdue University... Assim, eu fui subitamente envolvido na biologia.³

Por 10 anos, o campo da biologia ao qual Benzer aderiu envolveu a exaustiva análise de dois genes no cromossomo do bacteriófago T4. O estudo de Benzer desses genes produziu os mapas genéticos mais detalhados já construídos. Como veremos neste capítulo, também ajudou os geneticistas a esclarecer o que eles queriam dizer com a palavra “gene”. Assim, o livro de Schrödinger, envolvido com a questão “o que é a vida”, inspirou o trabalho que no final veio a responder a outra pergunta fundamental: “O que é um gene?”

O gene é para os geneticistas o que o átomo é para a química. Assim, ao longo deste texto enfocamos nossa atenção no gene e nas formas alternativas dos genes (alelos). Nos capítulos precedentes, examinamos os padrões de transmissão de genes que se distribuem independentemente e genes ligados, a localização cromossômica dos genes, a composição química de genes e cromossomos, o mecanismo de replicação, os eventos mutacionais nos genes e os mecanismos pelos quais os genes exercem seus efeitos sobre o fenótipo dos organismos. O que é esta unidade de informação genética que chamamos de gene? Como veremos, o conceito de gene não é estático. Ele se desenvolveu através de várias fases desde que Wilhelm Johannsen introduziu o termo em 1909, e sem dúvida evoluirá por meio de refinamentos adicionais no futuro.

O gene foi primeiro definido como a unidade de informação genética que controla um aspecto específico do fenótipo. Tal descrição, embora apurada, não nos dá uma definição precisa, não ambígua, que possa ser usada para identificar

um gene em termos moleculares. Em um nível mais fundamental, o gene foi definido como a unidade de informação genética que especifica a síntese de um polipeptídeo. Como o gene pode ser definido operacionalmente? Uma **definição operacional** seria uma operação ou experimento que se pode realizar para definir ou delimitar algo. Neste capítulo, enfocaremos o **teste de complementação** como uma definição operacional do gene, a unidade de informação genética que controla a síntese de um polipeptídeo ou uma molécula estrutural de RNA. Também consideraremos as limitações do teste de complementação, juntamente com as características estruturais únicas de genes selecionados.

EVOLUÇÃO DO CONCEITO DE GENE: RESUMO

Nosso conceito de gene sofreu muitos refinamentos desde a descoberta de Mendel dos “fatores unitários” em 1866.

Antes de discutirmos as evidências que apóiam os vários conceitos, vamos resumir os estágios importantes na

³Benzer, S. 1966. Adventures in the *rII* region. Em *Phage and the Origins of Molecular Biology* (J. Cairns, G. S. Stent e J. D. Watson, eds.) pp. 157-165. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

evolução do conceito de gene. A teoria gênica da herança começou com a publicação clássica de Mendel em 1866, mas não se tornou uma parte aceita do conhecimento científico até a redescoberta do trabalho de Mendel em 1900. O gene de Mendel (que não tinha este nome) era o “caráter” ou “fator unitário” que controlava uma característica fenotípica específica tal a cor da flor em ervilhas. Na época da redescoberta do trabalho de Mendel, o médico inglês Sir Archibald E. Garrod estava estudando várias características herdadas em humanos. Garrod reconheceu primeiro que a homozigose para alelos mutantes recessivos podia causar defeitos em processos normais do metabolismo. Seu conceito de gene é provavelmente dito de modo mais preciso como um gene mutante-um bloqueio metabólico, que cerca de 40 anos depois foi refinado para o conceito de um gene-uma enzima enunciado por George W. Beadle e Edward L. Tatum. Como muitas enzimas contêm dois ou mais polipeptídeos diferentes, cada um codificado por um gene separado, o conceito um gene-uma enzima subsequentemente foi

modificado para um gene-um polipeptídeo. Para casos em que o produto final de um gene é uma molécula de RNA, o conceito de gene pode ser modificado para um gene-um transcrito.

Antes de 1940, os genes foram considerados análogos a contas em um colar, e a recombinação ocorria entre genes mas não dentro deles. O gene era a unidade funcional básica, que controla uma característica fenotípica, e a unidade estrutural elementar, que não podia ser subdividida por recombinação ou mutação. O relato de Clarence Oliver de 1940 de que havia ocorrido recombinação dentro do gene *lozenge* de *Drosophila* estimulou muitos debates sobre seu significado. Quando o debate terminou, o par de nucleotídeos tinha substituído o gene como a unidade básica de estrutura, a unidade do material genético não subdivisível por mutação ou recombinação.

No início dos anos 1940, Edward B. Lewis desenvolveu o teste de complementação, ou *cis-trans*, para o alelismo funcional em *Drosophila*. Este teste foi subsequentemente

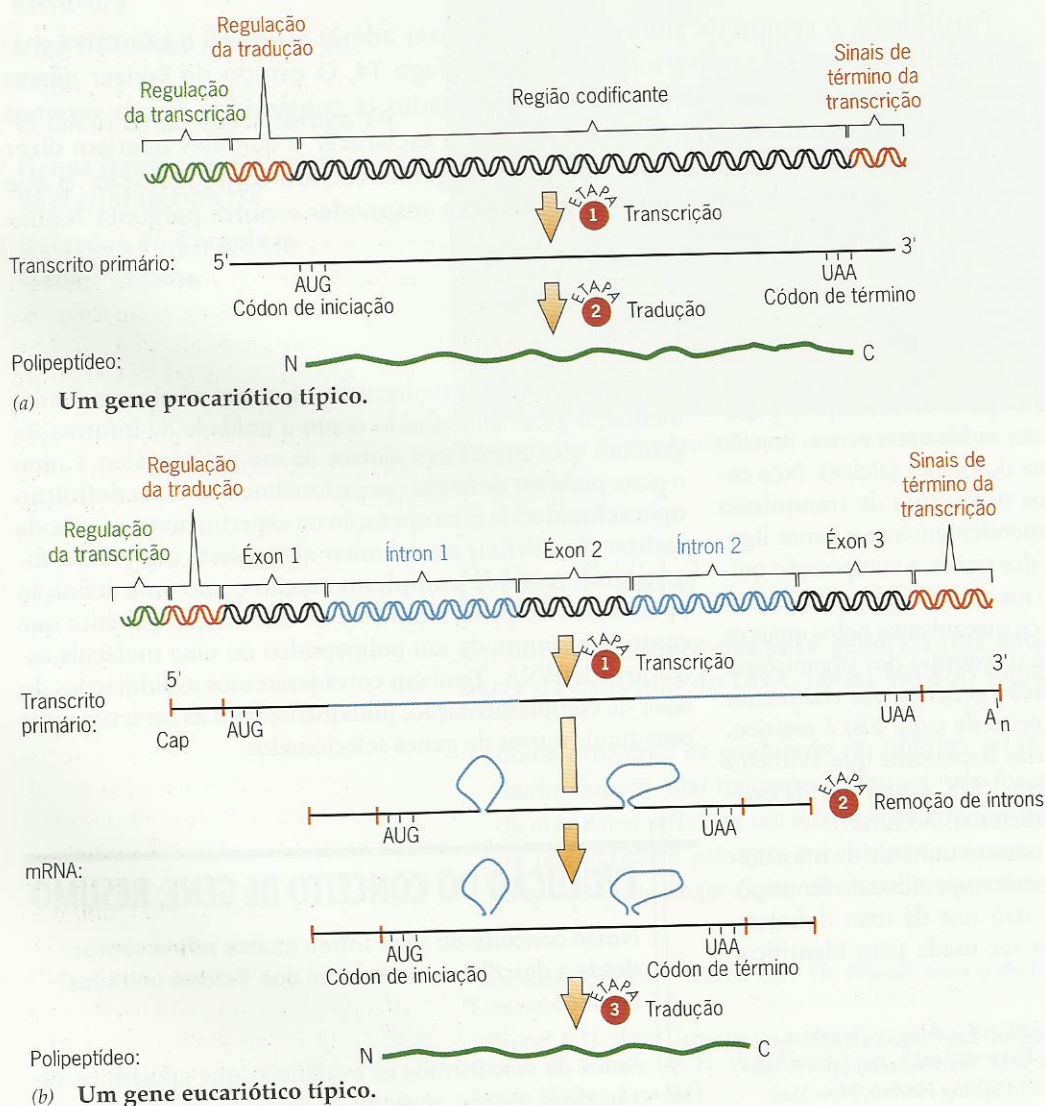


Fig. 14.1 ■ Estrutura gênica. Genes procarióticos em geral contêm seqüências codificantes ininterruptas (a), enquanto as seqüências codificantes de genes eucarióticos são comumente interrompidas por seqüências não codificantes (b).

explorado por Seymour Benzer para definir o gene experimentalmente no bacteriófago T4. O teste de complementação deu uma definição operacional do gene. Este teste permite a um geneticista determinar se duas mutações independentes estão situadas no mesmo gene ou em dois genes diferentes. Além disso, o gene, definido pelo teste de complementação, é um bom ajuste ao conceito um gene-um polipeptídeo.

Nos anos 1960, elegantes experimentos de Charles Yanofsky, Sydney Brenner e seus colaboradores mostraram que o gene e seu produto polipeptídico eram estruturas co-lineares, com uma correlação direta entre a sequência de pares de nucleotídeos no gene e a sequência de aminoácidos no polipeptídeo. Entretanto, este conceito simples de gene como uma sequência contínua de pares de nucleotídeos especificando uma sequência co-linear de aminoácidos no polipeptídeo produzido pelo gene teve vida curta. No final dos anos 1960, foram descobertos genes superpostos e genes dentro de genes, e foi demonstrado no final dos anos 1970 as sequências de genes eucarióticos que são interrompidas por sequências de íntrons. Além disso, mostrou-se que alguns genes, como por exemplo os genes codificantes de imunoglobulinas, são estocados em cromossomos da linhagem germinativa como “segmentos de genes”, que são reunidos em genes funcionais maduros durante o desenvolvimento.

Assim, a definição de gene precisa ser um tanto flexível para englobar todas as diferentes correlações estrutura-função que ocorrem nos organismos vivos. Aqui, definimos o gene como a unidade de informação genética que controla a síntese de um polipeptídeo ou uma molécula de RNA estrutural. Como definido, o gene pode ser identificado operacionalmente pelo teste de complementação. Como tal, o gene inclui as regiões 5' e 3' não codificantes que estão envolvidas na regulação da transcrição e tradução do gene e todos os íntrons dentro do gene (Fig. 14.1). O gene estrutural refere-se à parte que é transcrita para produzir o RNA. No caso de genes superpostos, esta definição requer que algumas sequências de pares de nucleotídeos sejam consideradas componentes de dois ou mais genes. Para os casos em que éxons são juntos em várias combinações para fazer proteínas correlatas mas diferentes, o gene pode ser definido como uma sequência de DNA que é uma unidade de transcrição e codifica um grupo de polipeptídeos proximamente correlatos, às vezes chamados “isoformas de proteínas”. Nos cromossomos da linhagem germinativa, as sequências de DNA que codificam segmentos de cadeias de anticorpos provavelmente devem ser chamadas “segmentos gênicos”, pois esta informação genética não está organizada em unidades que se ajustam a nenhuma das definições padrões de gene (Cap. 23).

■ PONTOS IMPORTANTES

- O conceito de gene sofreu muitos refinamentos desde o redescobrimto dos trabalhos de Mendel em 1900.
- A maioria dos genes codifica um polipeptídeo e pode ser definida operacionalmente pelo teste de complementação.

EVOLUÇÃO DO CONCEITO DE GENE: FUNÇÃO

Nosso conceito de gene evoluiu a partir do “fator unitário” de Mendel que controla uma característica fenotípica até a unidade do material genético que especifica um polipeptídeo e operacionalmente definido pelo teste de complementação.

Na seção anterior, examinamos resumidamente a evolução do conceito de gene, a unidade básica de informação genética. O resumo incluiu tanto o aspecto funcional quanto o estrutural do gene. Agora, vamos dar uma visão mais aprofundada do gene como uma unidade funcional.

Mendel: Fatores Constantes que Controlam Características Fenotípicas

A lei da combinação de caracteres diferentes, que controla o desenvolvimento dos híbridos, tem portanto sua fundamentação e exploração no princípio enunciado, de que os híbridos produzem ovócitos e grãos de pólen que em números iguais representam todas as formas constantes oriundas das combinações dos caracteres unidos na fertilização. (Mendel, 1866; tradução de William Bateson.)

Os caracteres ou fatores de Mendel, que hoje são denominados genes, controlam características fenotípicas específicas tais como cor das flores, cor das sementes e forma das sementes. Eles eram as unidades básicas do funcionamento, as unidades da informação genética que governa um aspecto específico do fenótipo. Esta definição de gene como unidade básica de função é aceita pela maioria dos geneticistas. Não houve mudança no conceito de gene como a unidade básica de função desde sua descoberta por Mendel em 1866. Entretanto, a elucidação da natureza química do material genético levantou dúvidas sobre a estrutura do gene, e nossa compreensão da estrutura molecular desta unidade básica de função sofreu vários refinamentos desde o redescobrimto do trabalho de Mendel em 1900.

Se examinarmos o que se sabe sobre como os genes controlam características fenotípicas, ficará óbvia a necessidade de uma definição mais precisa do gene. A via pela qual o gene exerce seu efeito sobre o fenótipo de um organismo geralmente é muito complexa (Fig. 14.2). Vários genes podem ter efeitos similares sobre a mesma característica fenotípica, dificultando separar os efeitos individuais dos genes. Todos os genes de um organismo estão situados nos mesmos núcleos, e eles não funcionam todos independentemente. O fenótipo de um organismo é o produto da ação de todos os genes dentro das restrições impostas pelo ambiente. Cada gene também tem um efeito sobre a população a qual pertence o organismo portador do gene. Finalmente, cada gene tem um efeito potencial, por pequeno que seja, sobre o fenótipo cumulativo da biosfera, pois cada gene pode afetar a capacidade do organismo, ou da população ou da espécie de competir por um nicho ecológico na biosfera (Cap. 27).

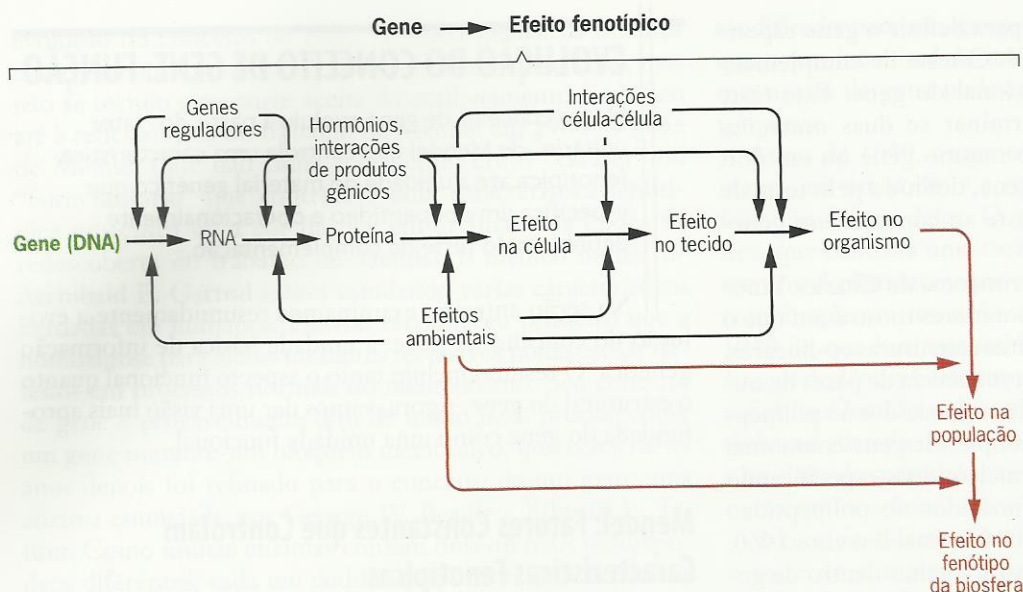


Fig. 14.2 ■ A via complexa pela qual um gene exerce seu efeito sobre o fenótipo de um organismo, uma população ou a biosfera.

Garrod: Um Gene Mutante-Um Bloqueio Metabólico

Na época do redescobrimento do trabalho de Mendel em 1900, Sir Archibald Garrod estava estudando várias doenças metabólicas congênitas em humanos. Uma delas era a doença herdada **alcaptonúria**, que é facilmente detectada pelo escurecimento da urina com exposição ao ar. A substância responsável pelo escurecimento é alcaptona (ou ácido homogentísico), um intermediário na degradação dos aminoácidos aromáticos tirosina e fenilalanina (Fig. 14.3). Garrod acreditava que a presença de ácido homogentísico na urina era devida a um bloqueio na via normal do metabolismo deste composto. Além disso, com base nos estudos de heredogramas, Garrod propôs que a alcaptonúria era herdada como um único gene recessivo. Os resultados dos estudos de Garrod sobre alcaptonúria e algumas outras doenças congênitas em humanos, tais como albinismo, foram apresentados em detalhes em seu livro *Erros Inatos do Metabolismo*. Embora os detalhes da via bioquímica afetada pelas mutações recessivas que causam alcaptonúria só tenham sido esclarecidos muitos anos depois, Garrod compreendeu claramente a relação entre genes e metabolismo. Seu conceito pode ser dito como **um gene mutante-um bloqueio metabólico**.

Beadle e Tatum: Um Gene-Uma Enzima

Durante o final dos anos 1930, George Beadle e Boris Ephrussi fizeram experimentos pioneiros sobre mutantes de cor de olho em *Drosophila*. Eles identificaram genes que são necessários para a síntese de pigmentos específicos dos olhos, indicando que vias metabólicas catalisadas por enzimas estão sob controle genético. Seus resultados motivaram Beadle a procurar um organismo ideal para usar na ampliação deste trabalho. Ele escolheu o mofo do pão com cor

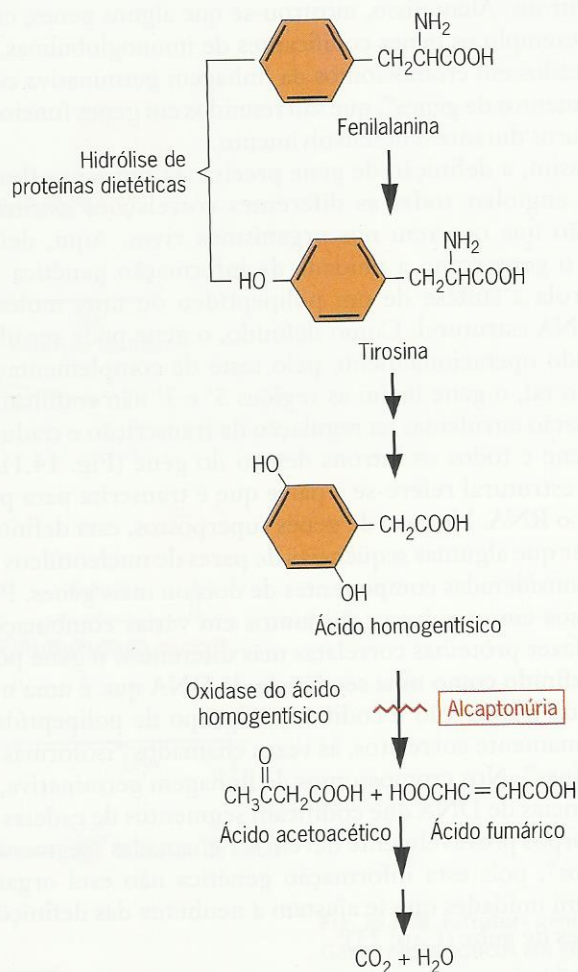


Fig. 14.3 ■ Alcaptonúria em humanos resulta de um bloqueio no catabolismo de fenilalanina causado por uma mutação no gene que codifica a enzima oxidase do ácido homogentísico. Quando esta enzima está ausente ou inativa, seu substrato, o ácido homogentísico, acumula-se nos tecidos e na urina.

de salmão *Neurospora crassa* porque ele pode crescer em um meio contendo apenas (1) sais inorgânicos, (2) um açúcar simples e (3) uma vitamina, a biotina. O meio de crescimento de *Neurospora* contendo apenas estes componentes é chamado "meio mínimo". Beadle e seu novo colaborador, Edward Tatum, raciocinaram que *Neurospora* deve ser capaz

de produzir *de novo* todos os outros metabólitos essenciais, tais como purinas, pirimidinas, aminoácidos e outras vitaminas. Além disso, imaginaram que a biossíntese destes fatores de crescimento deve estar sob controle genético. Sendo assim, esperava-se que mutações em genes cujos produtos estão envolvidos na biossíntese de metabólitos essenciais

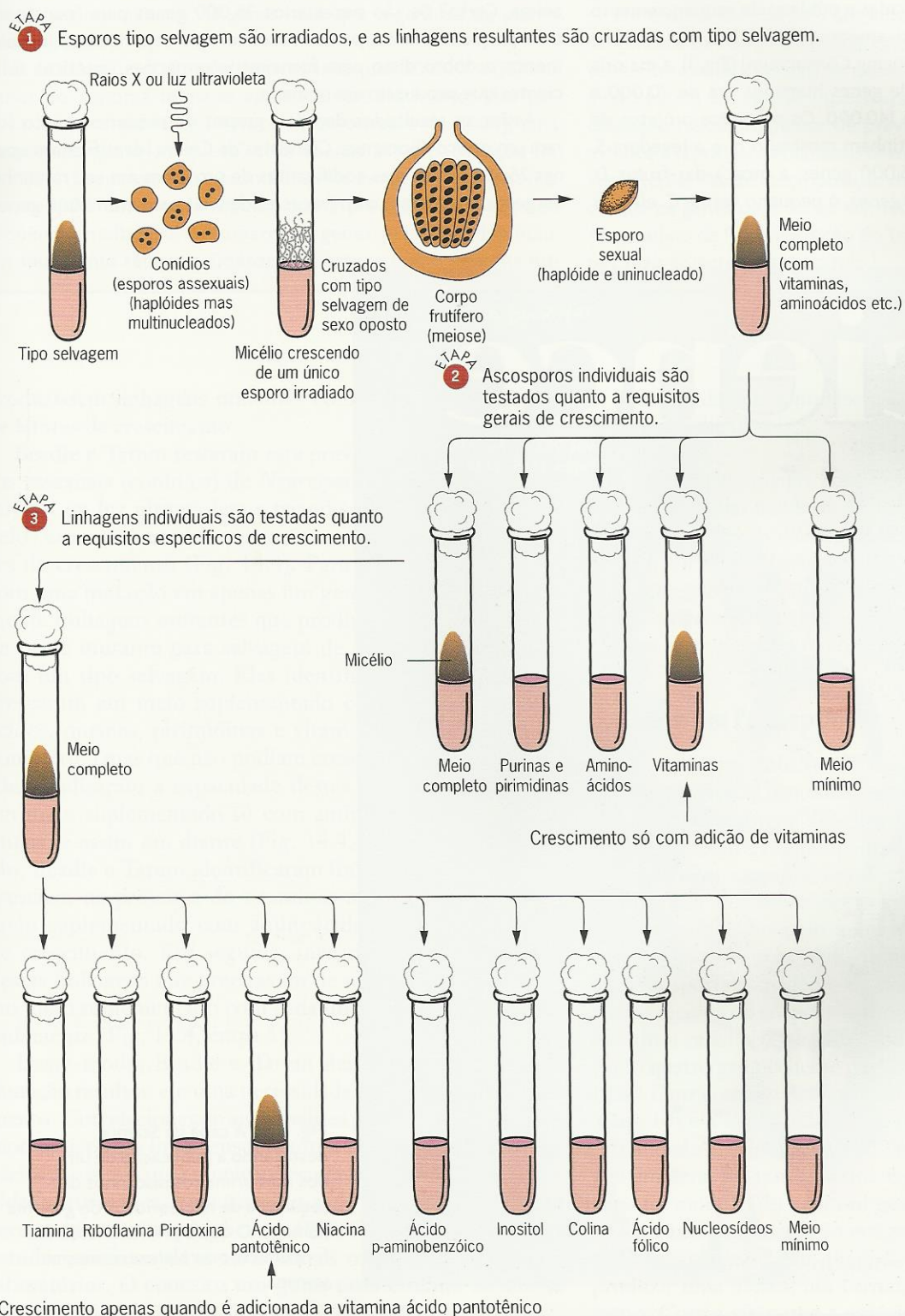


Fig. 14.4 ■ Diagrama do experimento de Beadle e Tatum com *Neurospora* que levou à hipótese um gene-uma enzima.



ENFOQUE TÉCNICO: O Genoma Humano – Quantos Genes?

Quanto genes são necessários para determinar o crescimento e o desenvolvimento de um ser humano? Antes da liberação dos primeiros rascunhos da sequência do genoma humano em fevereiro de 2001 por equipes privadas e públicas de sequenciamento do genoma humano (respectivamente, Celera, Inc. e o International Human Genome Sequencing Consortium) (Fig. 1), a maioria das estimativas do número de genes humanos era de 70.000 a 100.000, com algumas de até 140.000. Os primeiros projetos de sequenciamento do genoma tinham mostrado que a levedura *S. cerevisiae* possuía cerca de 6.000 genes, a mosca-das-frutas *D. melanogaster* cerca de 13.000 genes, o pequeno verme *C. elegans*

aproximadamente 19.000 genes e a pequena planta *A. thaliana* cerca de 26.000 genes. Certamente, nós humanos somos muito mais complexos do que qualquer uma das espécies mencionadas acima. Certo? Se são necessários 26.000 genes para “construir” uma pequena planta, parece razoável que seja necessário pelo menos o dobro disto para fornecer informações genéticas suficientes que produzam um humano.

Assim, os resultados dos dois grupos de sequenciamento foram um pouco chocantes. Cientistas da Celera identificaram apenas 26.588 seqüências codificantes de proteínas em seu rascunho do genoma humano, com fracas evidências para até 12.000 genes

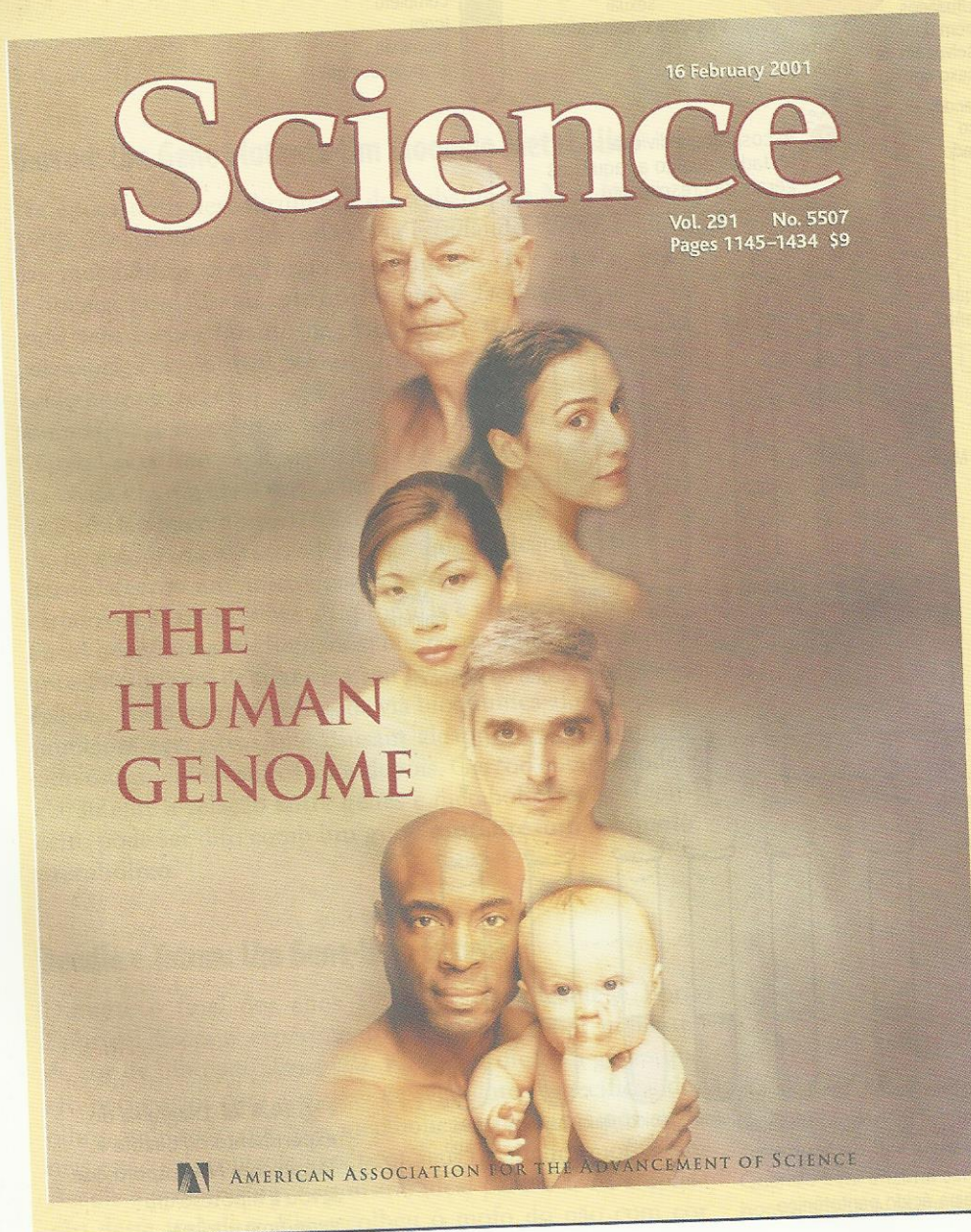


Fig. 1 ■ A capa da *Science* destacando a publicação de um dos dois primeiros rascunhos da sequência de nucleotídeos do genoma humano. O outro primeiro rascunho foi publicado na *Nature* na mesma semana.



ENFOQUE TÉCNICO: O Genoma Humano – Quantos Genes? (continuação)

adicionais. A análise do International Consortium de seus dados os levou a prever um total de 30.000 a 35.000 genes no genoma humano. Se eles estiverem corretos, o genoma humano contém apenas alguns milhares de genes mais que o genoma da planta *Arabidopsis*.

O International Consortium estima que 30.000 a 35.000 genes no genoma humano aparentemente é muito. Eles publicaram recentemente (*Nature*, 21 de outubro de 2004), uma “sequência quase completa”, com apenas 341 espaços, que inclui 99% da eucromatina (cromatina rica em genes) no genoma. Surpreendentemente, à medida que a precisão dos dados de sequência melhorou, o número de genes previstos diminuiu. Em seu último relato, o Consortium estima que o genoma hu-

mano contém apenas 20.000 a 25.000 genes que codificam proteínas. Talvez nós não sejamos tão complexos e sofisticados como pensávamos.

Entretanto, Eric Lander, um autor líder na publicação inicial do genoma humano do International Consortium, alerta que devemos ter cautela pois (1) os genes que se expressam em níveis muito baixos em células especializadas podem ter escapado da detecção pelos programas de previsão de genes usados nestes estudos e (2) os genes humanos dão origem a várias proteínas por vias alternativas de recomposição de éxons mais frequentemente do que os genes de moscas, vermes e plantas. (Veja a seção Vias Alternativas de Recomposição do Transcrito: Isoformas de Proteínas, mais adiante neste capítulo.)

produzissem linhagens mutantes com requisitos adicionais de fatores de crescimento.

Beadle e Tatum testaram esta previsão irradiando esporos assexuais (conídios) de *Neurospora* tipo selvagem com raios X ou luz ultravioleta e triando os clones produzidos pelos esporos mutados quanto a novos requisitos de fatores de crescimento (Fig. 14.4). Para selecionar linhagens com uma mutação em apenas um gene, eles estudaram somente linhagens mutantes que produziam uma proporção de prole mutante para selvagem de 1:1 quando cruzados com um tipo selvagem. Eles identificaram mutantes que cresceram em meio suplementado com todos os aminoácidos, purinas, pirimidinas e vitaminas (chamado “meio completo”) mas que não podiam crescer em meio mínimo. Eles analisaram a capacidade destes mutantes de crescer em meio suplementado só com aminoácidos, ou só vitaminas, e assim em diante (Fig. 14.4, etapa 2). Por exemplo, Beadle e Tatum identificaram linhagens mutantes que cresciam na presença de vitaminas mas não cresciam em meio suplementado com aminoácidos ou outros fatores de crescimento. Em seguida, investigaram a capacidade destas linhagens que precisavam de vitaminas para crescer em meio suplementado com cada uma das vitaminas separadamente (Fig. 14.4, etapa 3).

Deste modo, Beadle e Tatum demonstraram que cada mutação resultou em uma necessidade de um fator de crescimento. Correlacionando suas análises genéticas com estudos bioquímicos das linhagens mutantes, eles demonstraram em vários casos que uma mutação resultava na perda de uma atividade enzimática. Esse trabalho, pelo qual Beadle e Tatum receberam o Prêmio Nobel em 1958, foi logo verificado por estudos similares de muitos outros organismos em vários laboratórios. O conceito **um gene-uma enzima** tornou-se então um ponto central da genética molecular.

Apropriadamente, no discurso de recebimento do Prêmio Nobel, Beadle disse:

Neste longo caminho, descobrimos o que Garrod viu tão claramente há muitos anos. Hoje, conhecemos o seu trabalho e estamos cientes de que adicionamos pouco, se algo, em um princípio novo... Portanto, somos capazes de demonstrar que o que Garrod mostrou para alguns genes e algumas reações químicas em humanos era verdade para muitos genes e muitas reações em Neurospora.

Um Gene-Um Polipeptídeo

Subseqüente ao trabalho de Beadle e Tatum, soube-se que muitas enzimas e proteínas estruturais são heteromultiméricas, isto é, contêm duas ou mais cadeias polipeptídicas diferentes, com cada polipeptídeo codificado por um gene separado. Por exemplo, em *E. coli*, a enzima triptofano-sintetase é um heterotetrâmero composto de dois polipeptídeos α codificados pelo gene *trpA* e dois polipeptídeos β codificados pelo gene *trpB*. Similarmente, as hemoglobinas, que transportam oxigênio de nossos pulmões para todos os outros tecidos do corpo, são proteínas tetraméricas que contêm duas cadeias α -globina e duas cadeias β -globina, assim como quatro grupos heme para ligação de oxigênio (veja Fig. 12.3). Outras enzimas, como a DNA-polimerase III de *E. coli* (Cap. 10) e a RNA-polimerase II (Cap. 11), contêm muitas subunidades polipeptídicas diferentes, cada uma codificada por um gene separado. Assim, o conceito um gene-uma enzima foi modificado para **um gene-um polipeptídeo**.

Considerando a relação um gene-um polipeptídeo, os geneticistas perguntaram quantos genes são necessários para produzir uma mosca, um verme, uma planta ou um ser humano. Com a visão que a espécie humana tem de si mesma

como a mais avançada e altamente evoluída do planeta, a resposta a esta pergunta sobre o número de genes foi uma grande surpresa (veja Enfoque Técnico: O Genoma Humano – Quantos Genes?).

PONTOS IMPORTANTES

- A existência de um elemento genético básico, o gene, que controla uma característica fenotípica específica foi estabelecida pelo trabalho de Mendel em 1866.
- Durante o último século, o conceito de gene evoluiu da unidade que pode mutar com o objetivo de causar um bloqueio específico no metabolismo para a unidade que especifica uma enzima e para a sequência de pares de nucleotídeos no DNA que codificam uma cadeia polipeptídica ou uma molécula de RNA.

EVOLUÇÃO DO CONCEITO DE GENE: ESTRUTURA

Nosso conceito de gene evoluiu de “contas em um colar” não divisível por mutação ou recombinação para a sequência de pares de nucleotídeos que codifica um polipeptídeo ou, em alguns casos, uma molécula de RNA.

Na seção anterior, examinamos a evolução do conceito de gene como o componente funcional básico do material genético. Nesta seção, examinaremos o gene do ponto de vista estrutural. Qual a estrutura do gene? Todos os genes têm a mesma estrutura?

O Conceito Contas-de-um-Colar Anterior a 1940

Antes de 1940, os genes em um cromossomo eram considerados análogos a contas em um colar. Acreditava-se que a recombinação ocorria apenas entre as contas ou genes, e não dentro de genes. O gene parecia ser indivisível. De acordo com este conceito contas-em-um-colar, o gene era a unidade básica de informação genética definida por três critérios: (1) função, (2) recombinação e (3) mutação. Mais especificamente, o gene era

1. *A unidade de função*, a unidade do material genético que controlava a herança de uma “característica” ou um atributo do fenótipo.
2. *A unidade de estrutura*, operacionalmente definida de dois modos:
 - a. *Por recombinação*: como a unidade de informação genética não subdivisível por recombinação.
 - b. *Por mutação*: como a menor unidade do material genético capaz de mutação independente.

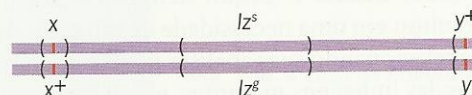
Os geneticistas inicialmente pensaram que todos os três critérios definiam a mesma unidade básica de herança, ou seja, o gene.

Os geneticistas hoje sabem que estes critérios definem duas unidades diferentes de herança. De acordo com o conceito molecular atual, o gene é a unidade de função, a unidade de informação genética que controla a síntese de uma cadeia polipeptídica ou, em alguns casos, uma molécula de RNA. A unidade de estrutura é simplesmente a unidade estrutural no DNA, o par de nucleotídeos. Como não faz sentido chamar cada par de nucleotídeos de gene, os geneticistas enfocaram a definição original de gene como a unidade de função e descartaram a visão contas-em-um-colar de que o gene não é subdivisível por recombinação ou mutação. Isto é claramente apropriado desde a ênfase no trabalho de Mendel sobre o fator (ou gene, como chamamos agora) que controla uma característica fenotípica.

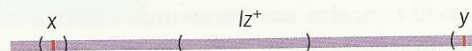
Descoberta de Recombinação Dentro do Gene

Em 1940, Clarence P. Oliver publicou a primeira evidência indicando que podia ocorrer recombinação dentro de um gene. Oliver estava estudando mutações no *locus lozenge* no cromossomo X de *Drosophila melanogaster*. Duas mutações, lz^s (olho de “óculos”) e lz^g (olho “vítreo”), foram consideradas alelos, isto é, formas diferentes do mesmo gene. Os dados disponíveis antes de 1940 indicavam que eles estavam mapeados no mesmo *locus* no cromossomo X. Eles tinham efeitos similares sobre o fenótipo do olho, e as fêmeas heterozigotas lz^s/lz^g tinham olhos *lozenge* em vez de olhos tipo selvagem. Entretanto, quando fêmeas lz^s/lz^g foram cruzadas com machos lz^s ou lz^g e foi observado um grande número de prole, a prole selvagem ocorria com frequência de cerca de 0,2 por cento.

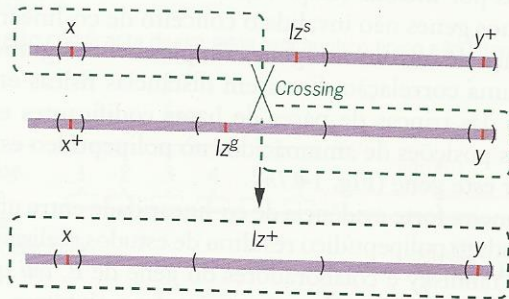
Essa rara prole tipo selvagem poderia ser explicada por reversão da mutação lz^s ou lz^g . Mas havia dois argumentos fortes contra a explicação de reversão. (1) A frequência de reversão de lz^s ou lz^g para tipo selvagem em machos *lozenge* hemizigotos foi muito menor que 0,2 por cento. (2) Quando as fêmeas heterozigotas lz^s/lz^g tinham marcadores genéticos delimitando o *locus lozenge*, a rara prole com olhos tipo selvagem sempre tinha um cromossomo X com lz^+ que era flanqueado por marcadores recombinantes externos. Além disso, a mesma combinação de marcadores externos sempre ocorria, como se os sítios lz^s e lz^g fossem fixos um em relação ao outro e ocorresse *crossing* entre eles. Se a fêmea heterozigota lz^s/lz^g tivesse um cromossomo X do tipo



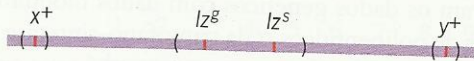
a rara prole com olhos tipo selvagem toda (com uma exceção) conteria um cromossomo X com a seguinte composição:



Entre a prole destes cruzamentos, a combinação recíproca de marcadores externos ($x^+ y^+$) nunca aparecia em combinação com lz^+ . Este resultado sugeriu fortemente que as mutações lz^s e lz^g estivessem situadas em pontos distantes no locus *lozenge* e que o cromossomo lz^+ era produzido por *crossing* entre os dois sítios, como mostrado no diagrama seguinte:



Evidências definitivas do envolvimento de recombinação exigiram a recuperação e a identificação do mutante duplo lz^s-lz^g com a combinação recíproca de marcadores externos, isto é,



Oliver não foi capaz de identificar este mutante duplo porque ele não poderia ser distinguido dos fenótipos parentais com uma só mutação. A identificação de ambos os produtos, os cromossomos tipo selvagem e mutantes duplos, produzidos por *crossing-over* dentro do gene *lozenge*, foi primeiro feita por Melvin M. Green, um dos estudantes de Oliver.

Os resultados destes estudos pioneiros primeiro indicaram que o gene era mais complexo do que uma conta em um colar. Eles mostraram que o gene era divisível, contendo sítios que eram separáveis por *crossing-over*. Os resultados de Oliver e Green foram a primeira etapa para o conceito atual de gene como uma longa sequência de pares de nucleotídeos, capazes de mutar e se recombinar em muitos sítios diferentes ao longo de seu tamanho.

Recombinação Entre Pares Adjacentes de Nucleotídeos

Os resultados obtidos por Oliver e Green em seus estudos de genes de *Drosophila* indicaram que sítios mutantes separáveis por recombinação podiam existir dentro de um único gene. Seymour Benzer ampliou este quadro do gene demonstrando a existência de 199 sítios distintos de mutação que eram separáveis por recombinação dentro do gene *rIIA* do bacteriófago T4 (veja O Locus *rII* do Fago T4). O conceito de gene de Benzer como uma sequência de pares de nucleotídeos capaz de mutar em muitos sítios distintos foi logo verificado pelos resultados de muitos pesquisadores que investigaram a estrutura do gene em vários organismos diferentes, tanto procariontes quanto eucariontes. Considerando esta informação sobre a estrutura dos genes e a estru-

tura conhecida do DNA, sabemos que a menor unidade do material genético capaz de mutação pode ser um único par de nucleotídeos e que pode ocorrer recombinação entre pares adjacentes de nucleotídeos, seja entre genes ou dentro deles. A recombinação entre pares adjacentes de nucleotídeos de um gene foi primeiro demonstrada por Charles Yanofsky em seus estudos do gene *trpA* que codifica o polipeptídeo α de triptofano-sintetase em *E. coli*. Esta enzima, um tetrâmero contendo dois polipeptídeos α e dois β , catalisa a etapa final na biossíntese do aminoácido triptofano.

Yanofsky e seus colegas isolaram e caracterizaram um grande número de auxotróficos de triptofano com mutações no gene *trpA*. O gene tipo selvagem *trpA* codifica um polipeptídeo α com 268 aminoácidos de tamanho. Yanofsky e associados usaram técnicas laboriosas de seqüenciamento de proteínas para determinar a seqüência completa de aminoácidos do polipeptídeo α tipo selvagem. Eles também determinaram as substituições de aminoácidos que ocorreram em várias formas mutantes do polipeptídeo α de triptofano-sintetase. Eles mapearam as mutações dentro do gene *trpA* por cruzamentos de dois e três fatores e compararam as posições de mapa com os locais de substituição de aminoácidos nos polipeptídeos mutantes.

Yanofsky e colaboradores mostraram que pode ocorrer recombinação entre mutações que alteram o mesmo aminoácido. As mutações *trpA23* e *trpA46* resultam ambas na substituição de outro aminoácido (arginina no caso de *A23*, ácido glutâmico no caso de *A46*) para a glicina presente na posição 211 do polipeptídeo α de triptofano-sintetase tipo selvagem. Entretanto, estas duas mutações ocorrem em sítios mutáveis diferentes; isto é, os sítios *A23* e *A46* são separáveis por recombinação. Yanofsky e colaboradores determinaram os aminoácidos presentes na posição 211 do polipeptídeo α em outros mutantes bem como em revertentes e revertentes parciais dos mutantes *trpA23* e *trpA46*. Usando esta informação e as atribuições conhecidas de códons, eles foram capazes de determinar quais dos códons de glicina, arginina e ácido glutâmico estavam presentes no mRNA de *trpA* na posição que codifica o aminoácido 211 dos polipeptídeos α presentes em células *trp^+*, *trpA23* e *trpA46*, respectivamente (Fig. 14.5).

Uma vez conhecidos os códons específicos no mRNA, as seqüências correspondentes de pares de bases no gene estrutural do qual é transcrito o mRNA são também conhecidas. Um filamento do DNA será complementar ao mRNA, e o segundo filamento do DNA será complementar ao primeiro filamento. Portanto, os dados de Yanofsky demonstraram que os eventos mutacionais que produziram os alelos *trpA23* e *trpA46* eram transições G-C para A-T em pares de nucleotídeos adjacentes. As células *trp^+* produzidas por recombinação entre cromossomos com as mutações *A23* e *A46* demonstraram que tinha ocorrido recombinação entre pares de nucleotídeos adjacentes, como mostrado na Fig. 14.6. Os resultados de Yanofsky mostraram que a unidade de material genético não divisível por recombinação é o único par de nucleotídeos.

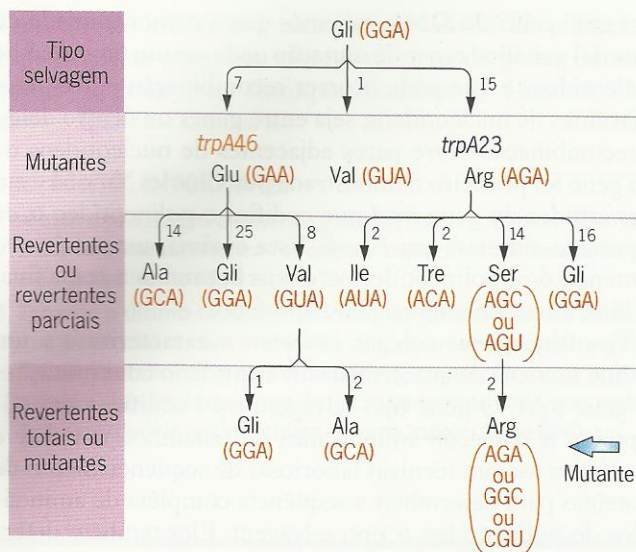


Fig. 14.5 ■ Sequência do aminoácido 211 (do terminal NH₂) do polipeptídeo α de triptofano sintetase de *E. coli*. O aminoácido 211 está alterado em mutantes *trpA23* e *trpA46* (veja Fig. 14.8). As trincas de códons mostradas entre parênteses são os únicos códons específicos para os aminoácidos indicados que irão permitir que todas as substituições observadas de aminoácidos ocorram por substituições de um único par de bases. O número ao lado de cada seta indica o número de vezes que foi observada a substituição em particular.

Co-linearidade Entre a Sequência Codificante de um Gene e seu Produto Polipeptídico

A informação genética é estocada em seqüências lineares de pares de nucleotídeos no DNA (ou nucleotídeos no RNA, em alguns casos). A transcrição e a tradução convertem esta informação genética em seqüências lineares de aminoácidos em polipeptídeos, que funcionam como os principais intermediários no controle genético do fenótipo.

Hoje sabemos que as seqüências de pares de nucleotídeos das regiões codificantes dos genes estruturais e as seqüências de aminoácidos dos polipeptídeos que eles codificam são

co-lineares: os primeiros três pares de bases da seqüência codificante de um gene especificam o primeiro aminoácido do polipeptídeo, os três pares de bases seguintes (quatro a seis) especificam o segundo aminoácido, e assim em diante, de um modo sistemático (Fig. 14.7a). Também é conhecido que a maioria dos genes em eucariontes superiores são interrompidos por íntrons (Cap. 11). Entretanto, a presença de íntrons nos genes não invalida o conceito de co-linearidade. A presença de íntrons nos genes simplesmente significa que não há uma correlação direta em distâncias físicas entre as posições das trincas de pares de bases codificantes em um gene e as posições de aminoácidos no polipeptídeo especificado por este gene (Fig. 14.7b).

A primeira forte evidência de co-linearidade entre um gene e seu produto polipeptídico resultou de estudos realizados por Charles Yanofsky e colaboradores do gene de *E. coli* que codifica a subunidade α da enzima triptofano-sintetase. Como já foi mencionado, esta enzima contém dois polipeptídeos α codificados pelo gene *trpA* e dois polipeptídeos β codificados pelo gene *trpB*. Yanofsky e colaboradores fizeram uma detalhada análise genética de mutações no gene *trpA* e correlacionaram os dados genéticos com dados bioquímicos das seqüências de polipeptídeos α da triptofano-sintetase tipo selvagem e mutante. Eles demonstraram que há uma correlação direta entre as posições de mapa de mutações no gene *trpA* e as posições das substituições de aminoácidos resultantes no polipeptídeo α da triptofano-sintetase (Fig. 14.8).

Quase ao mesmo tempo, Sydney Brenner e colaboradores demonstraram uma co-linearidade similar entre as posições de mutações no gene do bacteriófago T4 que codifica a principal proteína estrutural da cabeça do fago e as posições no polipeptídeo afetadas por estas mutações. Brenner e colaboradores estudaram as mutações *âmbar* (UAG de término de cadeia) e demonstraram uma correlação direta entre o tamanho do fragmento do polipeptídeo produzido e a posição da mutação dentro do gene.

Evidência definitiva de co-linearidade foi fornecida por comparações diretas entre as seqüências de nucleotídeos dos genes e as seqüências de aminoácidos de seus produtos polipeptídicos. Um dos primeiros casos em que a seqüência de

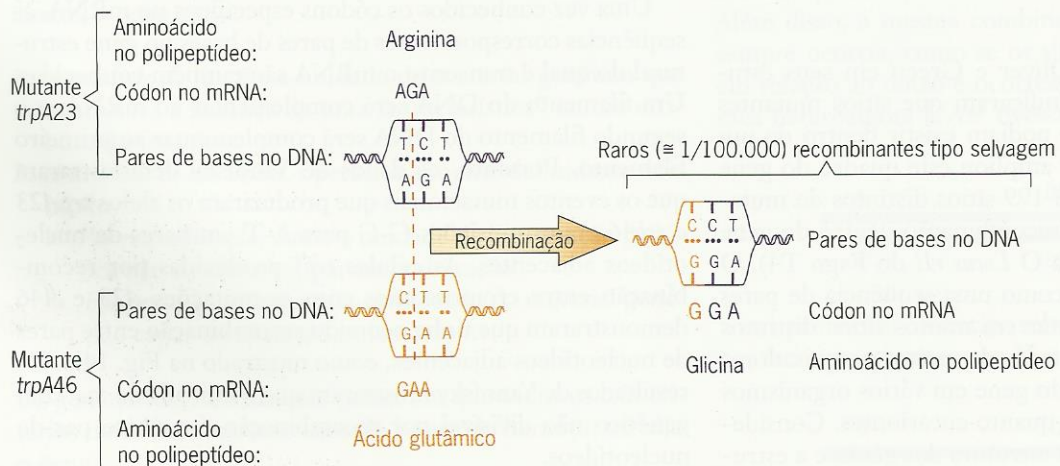
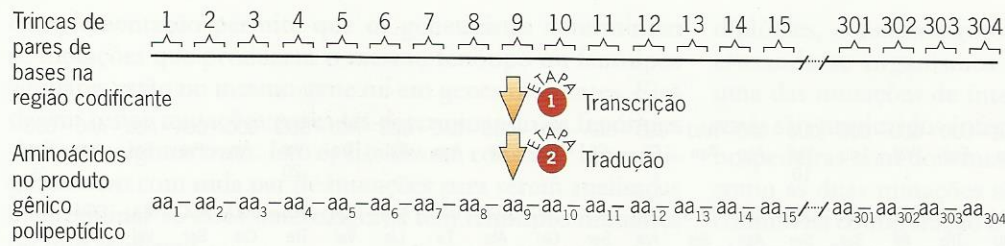
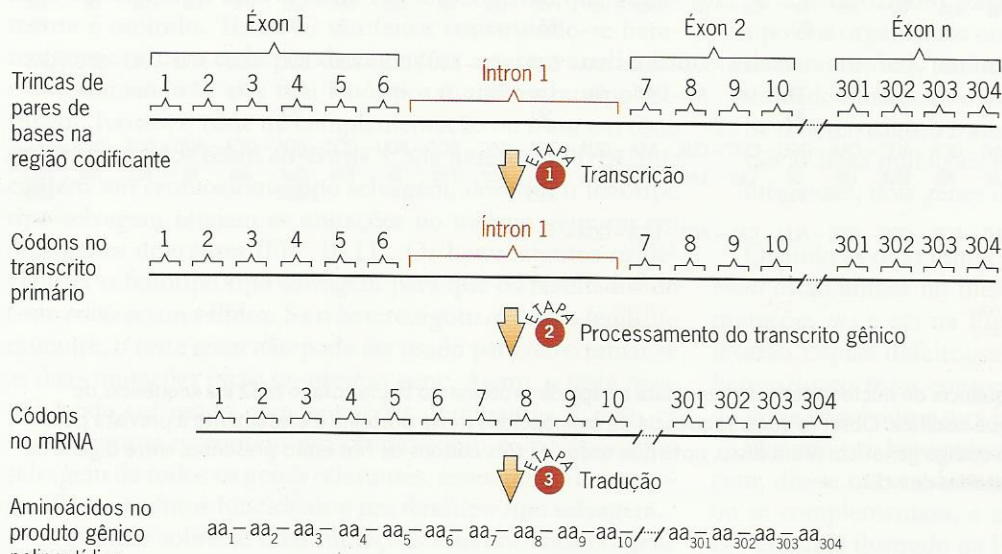


Fig. 14.6 ■ Recombinação entre mutações em pares adjacentes de nucleotídeos no gene *trpA* de *E. coli*. As mutações A23 e A46 resultam ambas em substituição de aminoácido na posição 211 do polipeptídeo de triptofano-sintetase α. A *E. coli* tipo selvagem tem uma glicina nesta posição do polipeptídeo α. A23 causa uma substituição de arginina por glicina; A46 causa uma substituição de ácido glutâmico por glicina (veja Fig. 14.5).



(a) Região codificante de um gene procariótico típico não interrompido.



(b) Região codificante de gene eucariótico típico interrompido por íntrons.

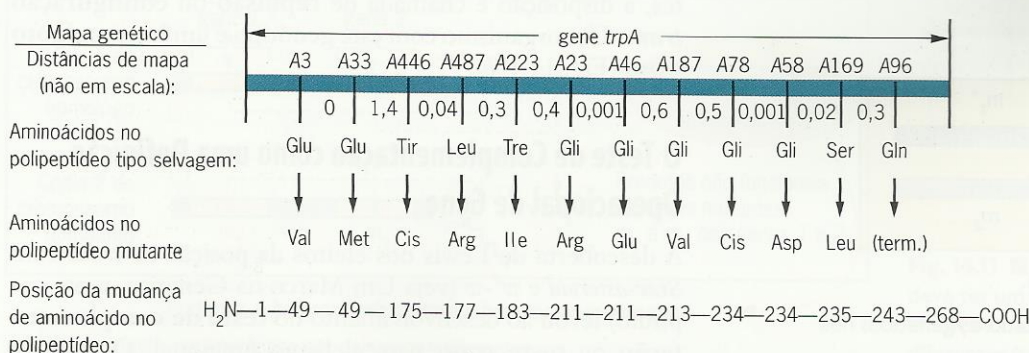
Fig. 14.7 ■ Co-linearidade entre as regiões codificantes de genes e seus produtos polipeptídicos.

aminoácidos de um polipeptídeo e a sequência de nucleotídeos do gene que o codifica foram ambas determinadas experimentalmente envolveu a proteína da capa do bacteriófago MS2 e o gene que a codifica. Este pequeno vírus tem um genoma de RNA que codifica apenas quatro proteínas, uma sendo a capa de proteínas que encapsula o RNA. Em 1972, Walter Fiers e colaboradores usaram o código genético para comparar a sequência de nucleotídeos do gene da capa de proteínas com a sequência de aminoácidos do polipeptídeo da capa e observaram que as sequências apresentavam uma co-linearidade perfeita (Fig. 14.9). Desde então, resultados

similares estabeleceram co-linearidade entre muitos genes e seus produtos protéicos em organismos que variam desde vírus até seres humanos.

PONTOS IMPORTANTES

- O conceito de gene evoluiu desde o de conta-em-um-colar, não divisível por recombinação ou mutação, até uma sequência de pares de nucleotídeos no DNA codificando uma cadeia polipeptídica.
- A unidade de material genético não divisível por recombinação ou mutação é um único par de nucleotídeos.

**Fig. 14.8** ■ Co-linearidade entre o gene *trpA* de *E. coli* e seu produto polipeptídico, o polipeptídeo α de triptofano-sintetase. As posições de mapa das mutações no gene *trpA* são mostradas em cima e as localizações das substituições de aminoácidos produzidas por estas mutações são mostradas abaixo do mapa.

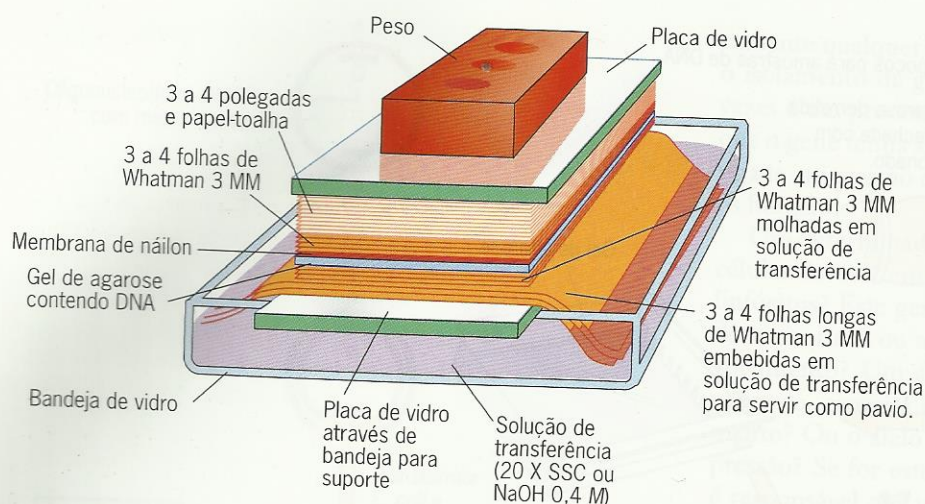


Fig. 15.22 ■ Procedimento usado para transferir DNA separados por eletroforese em gel para membranas de náilon. A solução de transferência leva o DNA do gel para a membrana à medida que as folhas de papel-toalha absorvem a solução salina do reservatório, através do gel, para as toalhas. O DNA liga-se à membrana por contato. A membrana com o DNA ligado a ela é seca e aquecida sob vácuo para afixar firmemente o DNA antes da hibridização. SSC é uma solução contendo cloreto de sódio e citrato de sódio.

posições das seqüências de DNA que se hibridizaram com a sonda (Fig. 15.23).

A capacidade de transferir moléculas de DNA que foram separadas por eletroforese em gel para membranas de náilon para estudos de hibridização e outros tipos de análises provou ser extremamente útil (veja Enfoque Técnico: Detecção de um Gene Mutante Causador de Fibrose Cística).

Análise de RNA por Hibridização Northern

Se moléculas de DNA podem ser transferidas de géis de agarose para membranas de náilon para estudos de hibridização, podemos esperar que moléculas de RNA separadas por eletroforese em gel de agarose possam ser similarmente transferidas e analisadas. De fato, tais transferências de RNA são usadas rotineiramente em laboratórios de genética. As transferências

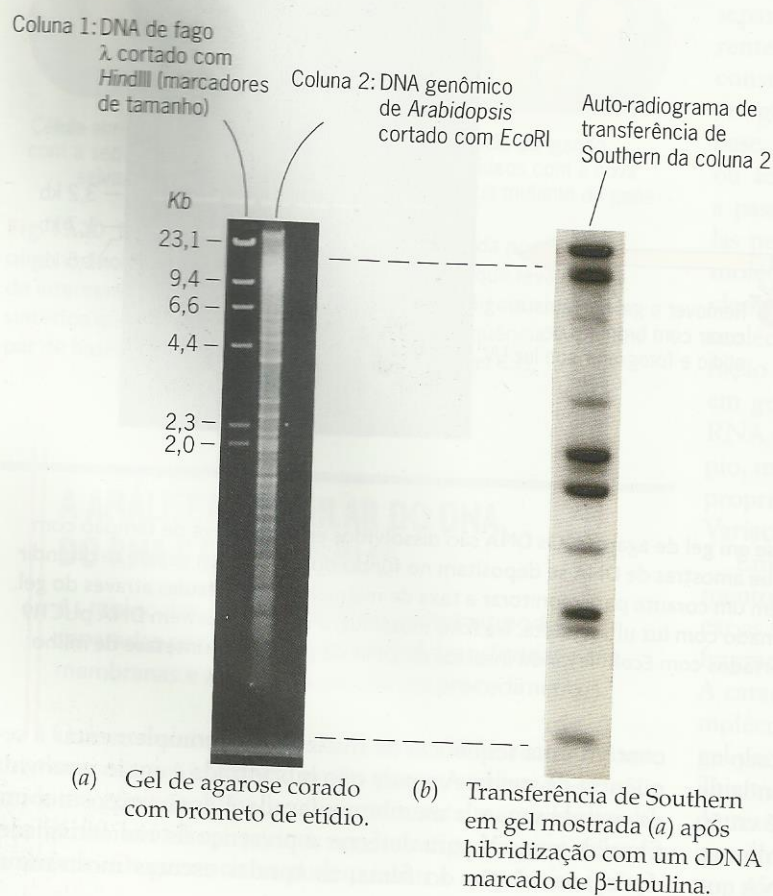


Fig. 15.23 ■ Identificação de fragmentos de restrição genômicos que possuem seqüências específicas de DNA pelo procedimento de hibridização de transferência de Southern. (a) Foto de um gel de agarose corado com brometo de etídio contendo DNA de fago λ digerido com HindIII (coluna da esquerda) e DNA de *Arabidopsis thaliana* digerido com EcoRI (coluna da direita). A digestão de DNA de λ fornece marcadores de tamanho. A digestão de DNA de *A. thaliana* foi transferida para uma membrana de náilon pelo procedimento de Southern (Fig. 15.22) e hibridizada com um fragmento de DNA radioativo de um gene clonado de β -tubulina. A transferência de Southern resultante é mostrada em (b); nove fragmentos diferentes de EcoRI hibridizados com a sonda de β -tubulina.

de RNA são chamadas **transferências northern**, uma brincadeira com o nome de Southern, em um procedimento análogo à transferência de Southern mas com moléculas de RNA sendo separadas e transferidas para uma membrana. Como será discutido na seção seguinte, esta terminologia foi ampliada para a transferência de proteínas de géis para membranas, um procedimento chamado de transferência *western*.*

O procedimento de transferência *northern* é essencialmente idêntico ao usado nas transferências de Southern (Fig. 15.22). Entretanto, as moléculas de RNA são muito sensíveis a degradação por RNases. Assim, deve ser tomado cuidado para evitar contaminação de materiais com estas enzimas extremamente estáveis. Além disso, a maioria das moléculas de RNA contém uma considerável estrutura secundária e devem portanto ser desnaturadas durante eletroforese de modo a separá-las com base no tamanho. A desnaturação é feita por adição de formaldeído ou algum outro desnaturante químico ao tampão de eletroforese usado. Após transferência para uma membrana apropriada, o RNA é hibridizado com sondas de RNA ou DNA como na transferência de Southern.

As hibridizações *northern* (Fig. 15.24) são extremamente úteis em estudos de expressão gênica. Elas podem ser usa-

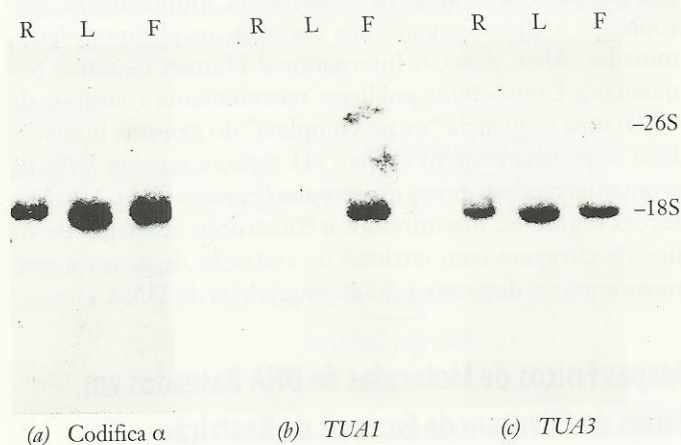


Fig. 15.24 ■ Dados típicos de hibridização *northern*. Os RNA totais foram isolados de raízes (R), folhas (L) e flores (F) de uma planta *A. thaliana*, separados por eletroforese em gel de agarose e então transferidos para membranas de náilon. O autoradiograma mostrado em (a) é uma transferência que foi hibridizada com uma sonda radioativa contendo uma sequência codificante de α -tubulina. Esta sonda hibridiza com os transcritos de todos os seis genes de α -tubulina em *A. thaliana*. Os autoradiogramas mostrados em (b) e (c) são de RNA que foram hibridizados com sondas de DNA específicas para os genes $\alpha 1$ e $\alpha 3$ de tubulina (*TUA1* e *TUA3*, respectivamente). Os resultados mostram que o transcrito de tubulina $\alpha 3$ está presente em todos os órgãos analisados, enquanto o transcrito de tubulina $\alpha 1$ só está presente nas flores. Os RNA ribossômicos 18S e 26S fornecem marcadores de tamanho. Suas posições foram determinadas a partir de uma foto do gel corado com brometo de etídio antes da transferência dos RNA para a membrana de náilon.

*N.T.: Northern e western são duas “brincadeiras” com o nome de Edwin Southern.

das para determinar quando e onde um determinado gene é expresso. Entretanto, devemos lembrar que as hibridizações *northern* só medem o acúmulo de RNA transcritos. Elas não dão informações sobre por que ocorreu o acúmulo observado. As mudanças nos níveis dos transcritos podem ser devidas a mudanças na taxa de transcrição ou a mudanças na taxa de renovação do transcrito. Devem ser usados procedimentos mais sofisticados para distinguir entre estas possibilidades.

Análise de Proteínas por Técnicas Western

A eletroforese em gel de poliacrilamida é um instrumento importante para separação e caracterização de proteínas. Como muitas proteínas funcionais são compostas de duas ou mais subunidades, polipeptídeos individuais são separados por eletroforese na presença do detergente sódico sulfato de dodecila (SDS), que desnatura proteínas. Após eletroforese, as proteínas são detectadas por coloração de azul Coomassie ou coloração de prata. Entretanto, os polipeptídeos separados também podem ser transferidos do gel para uma membrana de nitrocelulose, e proteínas individuais podem ser detectadas com anticorpos. Esta transferência de proteínas de géis de acrilamida para membranas de nitrocelulose, chamada de **transferência western**, é feita usando-se uma corrente elétrica para mover as proteínas do gel para a superfície da membrana (Fig. 15.25). Após a transferência, uma proteína específica de interesse é identificada colocando-se

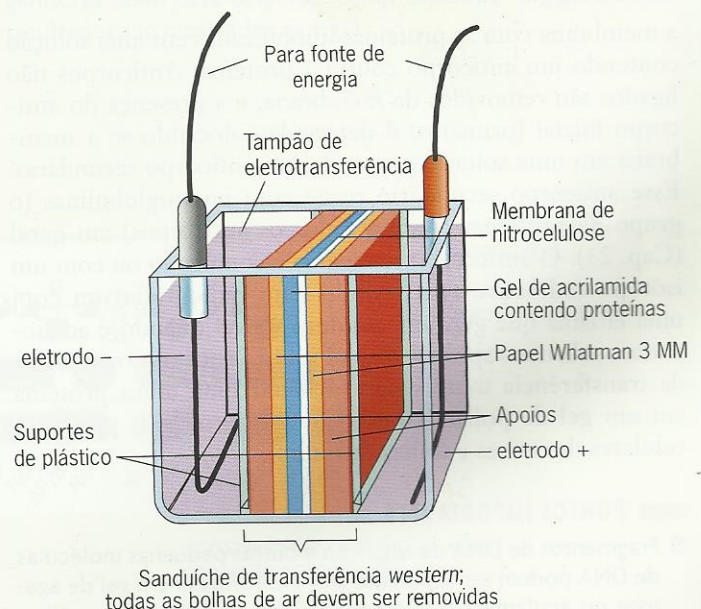


Fig. 15.25 ■ Uma típica transferência *western*, ou eletrotransferência. É usada uma corrente elétrica para transferir as proteínas de um gel para uma membrana de nitrocelulose colocada próxima a ele na transferência sanduiche. Todos os outros componentes do sanduiche funcionam para dar um suporte suave mas firme. Contato íntimo entre o gel e a membrana é essencial para boa transferência.

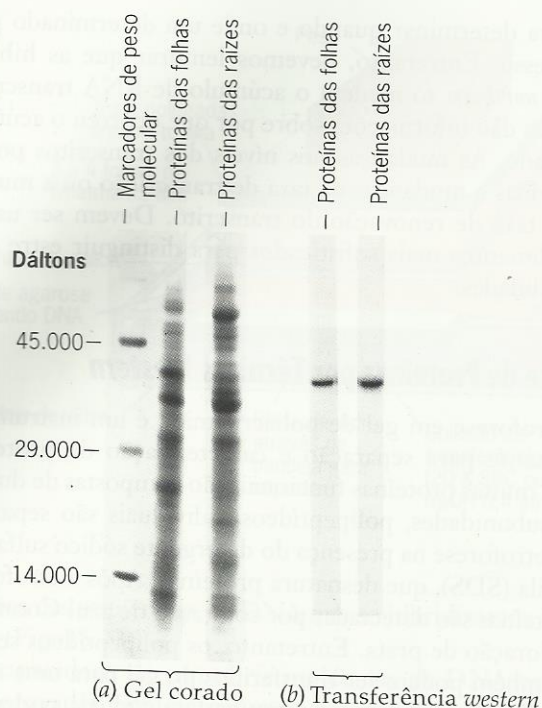


Fig. 15.26 ■ O uso de transferências *western* para identificar proteínas individuais separadas por eletroforese em gel de poliácridamida. (a) Proteínas isoladas de raízes ou folhas de milho foram separadas por eletroforese em gel de poliácridamida e coradas com Coomassie azul. (b) A forma cloroplástica de glutamina-sintetase foi identificada por análise de transferência *western* do gel mostrado em (a).

a membrana com as proteínas imobilizadas em uma solução contendo um anticorpo contra a proteína. Anticorpos não ligados são removidos da membrana, e a presença do anticorpo inicial (primário) é detectada colocando-se a membrana em uma solução contendo um anticorpo secundário. Esse anticorpo secundário reage com imunoglobulinas (o grupo de proteínas contém todos os anticorpos) em geral (Cap. 23). O anticorpo secundário é conjugado ou com um isótopo radioativo (permitindo auto-radiografia) ou com uma enzima que gera um produto visível quando é adicionado o substrato apropriado. A Fig. 15.26 mostra o emprego da transferência *western* para detectar uma única proteína em um gel de poliácridamida contendo todas as proteínas celulares das raízes e folhas de milho.

PONTOS IMPORTANTES

- Fragmentos de DNA de restrição e outras pequenas moléculas de DNA podem ser separados por eletroforese em gel de agarose ou acrilamida e transferidos para membranas de náilon para produzir transferências de Southern.
- Os DNA em transferências de Southern podem ser hibridizados com sondas marcadas de DNA para detectar seqüências de interesse por auto-radiografia.
- Quando moléculas de RNA são separadas por eletroforese em gel e transferidas para membranas para análise, os RNA resultantes no gel são denominados transferências *northern*.

- Quando proteínas são transferidas de géis para membranas e detectadas com anticorpos, os produtos são denominados transferências *western*.

A ANÁLISE MOLECULAR DE GENES E CROMOSSOMOS

Os sítios nos quais enzimas de restrição cortam moléculas de DNA podem ser usados para construir mapas físicos das moléculas. Entretanto, seqüências de nucleotídeos fornecem os mapas físicos finais de moléculas de DNA.

Técnicas de DNA recombinante permitem que geneticistas determinem a estrutura de genes, cromossomos e genomas inteiros. Pesquisadores de genética molecular estão construindo mapas genéticos e físicos detalhados dos genomas de muitos organismos (Cap. 16).

O mapa físico de um elemento genético é a sua seqüência de nucleotídeos, e as seqüências completas de nucleotídeos dos genomas de muitos vírus, bactérias, mitocôndrias, cloroplastos e alguns organismos eucarióticos já foram determinadas. Além disso, o International Human Genome Sequencing Consortium publicou recentemente (outubro de 2004) uma seqüência "quase completa" do genoma humano. Essa seqüência contém apenas 341 espaços e cobre 99% da cromatina rica em genes no genoma humano (Cap. 16). Nas seções seguintes, discutiremos a construção de mapas de sítios de clivagem com enzimas de restrição de genes e cromossomos e a determinação de seqüências de DNA.

Mapas Físicos de Moléculas de DNA Baseados em Sítios de Clivagem de Enzimas de Restrição

A maioria das endonucleases de restrição corta moléculas de DNA de um modo sítio-específico (veja Quadro 15.1). Como resultado, elas podem ser usadas para gerar **mapas físicos** de cromossomos que são de grande valor para ajudar pesquisadores no isolamento de fragmentos de DNA portadores de genes ou outras seqüências de DNA de interesse. Os tamanhos dos fragmentos de restrição podem ser determinados por eletroforese em gel de poliácridamida ou agarose (Fig. 15.21). Devido à subunidade estrutural do DNA em nucleotídeos, com um grupo fosfato por nucleotídeo, o DNA tem uma carga essencialmente constante por unidade de massa. Assim, as taxas de migração de fragmentos de DNA durante eletroforese fornecem estimativas precisas de seus tamanhos, sendo a taxa de migração inversamente proporcional ao tamanho.

O procedimento que é usado para mapear os sítios de clivagem por enzimas de restrição é ilustrado na Fig. 15.27. Os tamanhos dos fragmentos de restrição são estimados



ENFOQUE TÉCNICO: Detecção de um Gene Mutante Causador de Fibrose Cística

A fibrose cística é caracterizada pelo acúmulo de muco nos pulmões, no pâncreas e no fígado com subsequente funcionamento defeituoso destes órgãos. É a doença hereditária mais comum em humanos descendentes de europeus do norte. No Cap. 17, discutiremos fibrose cística e a identificação e caracterização do gene que a causa. Aqui, enfocaremos o uso de PCR para amplificar os alelos *CF* no DNA genômico de membros de famílias afetadas pela doença e a detecção do alelo mutante mais comum por hibridização de Southern para marcar sondas oligonucleotídicas.

Aproximadamente 70% dos casos de *CF* resultam de um alelo mutante específico do gene *CF*. Este alelo mutante, *CF*ΔF508, contém uma deleção de três bases que elimina uma fenilalanina na posição 508 no produto polipeptídico. Como a sequência de nucleotídeos do gene *CF* é conhecida e como o alelo *CF*ΔF508 difere do alelo tipo selvagem em três pares de bases, é possível criar sondas oligonucleotídicas que se hibridizam especificamente com o alelo *CF* tipo selvagem ou com o alelo ΔF508 em condições apropriadas.

O gene *CF* tipo selvagem e o produto gênico têm as seguintes sequências de nucleotídeos e aminoácidos na região alterada pela mutação ΔF508 (adiante):

Gene *CF*⁺

bases no filamento

codificante: 5'-AAA GAA AAT ATC ATC TTT GGT GTT-3'

aminoácidos no

produto: NH₂-Lis Glu Asn Ile Ile Fen Gli Val-COOH

aminoácido 508

deletada em ΔF508

Enquanto o alelo ΔF508 e o produto têm estas sequências:

bases na região

codificante: 5'-AAA GAA AAT ATC AT...T GGT GTT-3'

deleção

aminoácidos

no produto: NH₂-Lis Glu Asn Ile Ile — Gli Val-COOH

Fen ausente

Com base nestas sequências de nucleotídeos, Lap-Chee Tsui e colaboradores sintetizaram oligonucleotídeos que envolvem esta região dos alelos mutante e tipo selvagem do gene *CF* e testaram sua especificidade. Eles demonstraram que, a 37°C sob um conjunto padrão de condições, uma sonda de oligonucleotídeos (oligo-N: 3'-CTTTTATAGTAGAAACCAC-5') só se hibridizava com o alelo tipo selvagem, enquanto outra (oligo-ΔF: 3'-TTCTTTTATAGTA...ACCACAA-5') hibridizava apenas com o alelo ΔF508. Seus resultados mostraram que a sonda oligo-ΔF podia ser usada para detectar o alelo ΔF508 no estado homozigoto ou heterozigoto. Quando Tsui e colaboradores usaram estas sondas oligonucleotídicas alelo-específicas para analisar pacientes com *CF* e seus genitores quanto à presença da mutação ΔF508, eles descobriram que muitos dos pacientes eram homozigotos para esta mutação, enquanto a maioria de seus genitores eram heterozigotos, como esperado. Alguns de seus resultados são mostrados na Fig. 1.

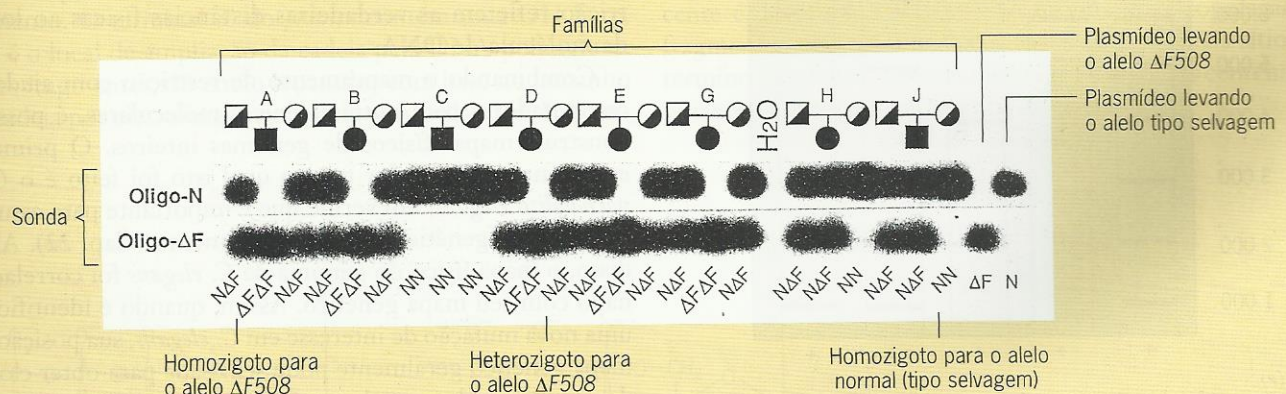


Fig. 1 ■ Detecção dos alelos tipo selvagem *CF* e ΔF508 por hibridização de sondas oligonucleotídicas aleloespecíficas marcadas com DNA genômico transferido para membranas de náilon pelo procedimento de transferência de Southern (Fig. 15.22). PCR foi usada para amplificar os loci *CF* nos DNA genômicos isolados de alguns membros familiares. Os produtos de PCR foram separados por eletroforese em gel, transferidos para membranas, desnaturados e hibridizados com sondas oligonucleotídicas radioativas (descritas acima). Foram preparadas transferências de Southern duplicadas; uma delas foi hibridizada com a sonda específica para o alelo *CF* tipo selvagem (fileira de cima) e a outra foi hibridizada com a sonda específica para o alelo ΔF508 (fileira de baixo). Os heredogramas das famílias mostrados em cima representam a prole com *CF* e seus genitores heterozigotos. Note que o alelo ΔF508 está presente nas famílias A, B, D, E e G. A família C tem um alelo *CF* diferente e as famílias H e J têm um genitor com o alelo ΔF508 e o outro genitor com um alelo *CF* diferente.

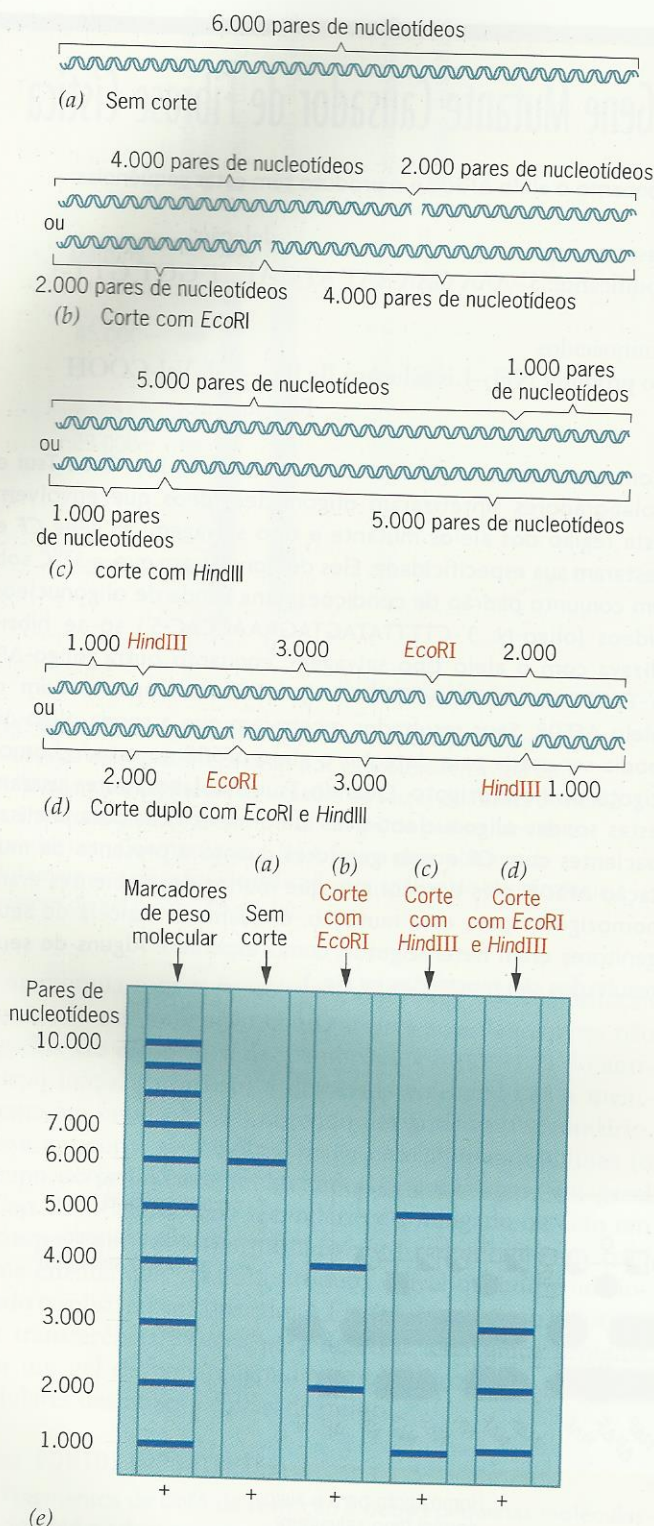


Fig. 15.27 ■ Procedimento usado para mapear sítios de clivagem com enzimas de restrição em moléculas de DNA. (a-d) Estruturas da molécula de DNA ou de fragmentos de restrição da molécula (a) não cortada ou cortada com (b) *EcoRI*, (c) *HindIII* ou (d) *EcoRI* e *HindIII*. (e) Separação destas moléculas de DNA e fragmentos por eletroforese em gel de agarose. A coluna da esquerda no gel contém um conjunto de marcadores de tamanho molecular, um conjunto de moléculas de DNA com 1.000 pares de nucleotídeos de tamanho e seus múltiplos.

usando-se um conjunto de marcadores de DNA de tamanhos conhecidos. Na Fig. 15.27, um conjunto de moléculas de DNA que diferem de tamanho em 1.000 pares de nucleotídeos é usado como marcador de tamanho. Considere uma molécula de DNA com aproximadamente 6.000 pares de nucleotídeos (6 kb) de tamanho. Quando a molécula de DNA de 6 kb é cortada com *EcoRI*, são produzidos dois fragmentos com os tamanhos de 4.000 e 2.000 pares de nucleotídeos. As possíveis posições do único sítio de corte por *EcoRI* na molécula são mostradas na Fig. 15.27b. Quando a mesma molécula de DNA é clivada com *HindIII*, resultam dois fragmentos com os tamanhos de 5.000 e 1.000 pares de nucleotídeos.

Os possíveis locais do único sítio de clivagem de *HindIII* são mostrados na Fig. 15.27c. Note que neste estágio da análise não se pode fazer nenhuma dedução sobre as posições relativas dos sítios de clivagem de *EcoRI* e *HindIII*. O sítio de clivagem de *HindIII* pode estar situado em um dos dois fragmentos de restrição de *EcoRI*. A molécula é então digerida simultaneamente com *EcoRI* e *HindIII*, e são produzidos três fragmentos com os tamanhos de 3.000, 2.000 e 1.000 pares de nucleotídeos. Este resultado estabelece as posições de dois sítios de clivagem em relação ao outro na molécula. Como o fragmento de restrição de *EcoRI* com 2.000 pares de nucleotídeos ainda está presente (não cortado por *HindIII*), o sítio de clivagem de *HindIII* deve estar na extremidade oposta ao sítio de clivagem de *EcoRI* na molécula (Fig. 15.27d). Estendendo este tipo de análise para incluir o uso de várias enzimas de restrição diferentes, podem ser construídos mapas mais amplos de sítios de restrição. Quando são empregados grandes números de enzimas de restrição, podem ser construídos mapas detalhados de cromossomos inteiros. Um aspecto importante dos **mapas de restrição** é que, ao contrário dos mapas genéticos (Cap. 7), os mapas de restrição refletem as verdadeiras distâncias físicas ao longo da molécula de DNA.

Combinando o mapeamento de restrição com ajuda de computadores com outras técnicas moleculares, é possível construir mapas físicos de genomas inteiros. O primeiro eucarionte multicelular para o qual isto foi feito é o *Caenorhabditis elegans*, um verme que é importante para estudos do controle genético do desenvolvimento (Cap. 22). Além disso, o mapa físico do genoma de *C. elegans* foi correlacionado com seu mapa genético. Assim, quando é identificada uma nova mutação de interesse em *C. elegans*, sua posição no mapa genético geralmente pode ser usada para obter clones do gene tipo selvagem de um grande banco internacional de clones de *C. elegans*.

Sequências de Nucleotídeos: Os Recentes Mapas de Ultra-estrutura

O mapa final da ultra-estrutura de um gene específico, ou cromossomo, é sua sequência de pares de nucleotídeos,

completa com a localização de todas as mudanças de pares de nucleotídeos que alteram a função deste gene ou cromossomo. Antes de 1975, a idéia de tentar sequenciar cromossomos inteiros era pouco provável, uma tarefa trabalhosa que iria demandar anos de trabalho. Entretanto, no final de 1976, todos os 5.386 nucleotídeos do cromossomo do fago ϕ X174 haviam sido sequenciados. Hoje, o sequenciamento é um procedimento laboratorial rotineiro. Sequências completas, ou quase completas, de nucleotídeos dos genomas de centenas de vírus, mais de 200 plasmídeos, cerca de 200 cloroplastos e mitocôndrias, mais de 200 bactérias e *archaea* e cerca de 30 eucariontes já são conhecidas. Além disso, a sequência de 99% da eucromatina no genoma humano já é conhecida (Cap. 16).

Nossa capacidade inicial de sequenciar praticamente qualquer molécula de DNA foi o resultado de quatro importantes desenvolvimentos. A descoberta mais importante foi a das enzimas de restrição e seu emprego na preparação de amostras homogêneas de segmentos específicos de cromossomos. Outro avanço importante foi a melhora dos procedimentos de eletroforese em gel ao ponto em que fragmentos de DNA que diferem em tamanho por um único nucleotídeo podem ser diferenciados. Também foram importantes as técnicas de clonagem gênica para facilitar a preparação de grandes quantidades de determinada molécula de DNA. Finalmente, pesquisadores inventaram dois procedimentos diferentes pelos quais as sequências de nucleotídeos de moléculas de DNA podiam ser determinadas.

Os procedimentos de sequenciamento de DNA dependem da geração de uma população de fragmentos de DNA todos tendo uma ponta em comum (todos terminam exatamente no mesmo nucleotídeo) e terminando em todas as posições possíveis (cada nucleotídeo consecutivo) na outra ponta. A ponta comum é o terminal 5' do *primer* de sequenciamento. O terminal 3' do *primer* contém uma OH livre, que é o local de ampliação da cadeia pela DNA-polimerase. Esses fragmentos são então separados com base no tamanho da cadeia por eletroforese em gel de poliácridamida. Em am-

bos os casos, são feitas quatro reações bioquímicas separadas simultaneamente, cada uma das quais gera um conjunto de fragmentos terminando em uma das quatro bases (A, G, C ou T) no DNA.

O primeiro procedimento, chamado procedimento de Maxam e Gilbert em homenagem a Allan Maxam e Walter Gilbert que o inventaram, usa quatro reações químicas diferentes para cortar cadeias de DNA especificamente em A, G, C ou C + T. O segundo procedimento, desenvolvido por Fred Sanger e colaboradores, usa a síntese de DNA *in vitro* na presença de nucleotídeos radioativos e finalizadores específicos de cadeia para gerar quatro populações de fragmentos radioativamente marcados que terminam em A, G, C e T, respectivamente. Discutiremos o procedimento de Sanger.

2',3'-Didesoxirribonucleosídeo-trifosfatos (Fig. 15.28) são os finalizadores de cadeia mais freqüentemente usados no procedimento de Sanger. Lembre que as DNA-polimerases têm necessidade absoluta de uma ponta 3'-OH livre no filamento *primer* de DNA (Cap. 10). Se for adicionado um 2',3'-didesoxinucleotídeo ao final de uma cadeia, ele irá bloquear a subsequente ampliação desta cadeia, pois os 2',3'-didesoxinucleotídeos não têm 3'-OH. Usando-se (1) 2',3'-didesoxitimidina-trifosfato (ddTTP), (2) 2',3'-didesoxicitidina-trifosfato (ddCTP), (3) 2',3'-didesoxiadenosina-trifosfato (ddATP) e (4) 2',3'-didesoxiguanosina-trifosfato (ddGTP) como finalizadores de cadeia em quatro reações separadas de síntese de DNA, podem ser geradas quatro populações de fragmentos, e cada população irá conter cadeias que terminam todas na mesma base (T, C, A ou G) (Fig. 15.29).

Em uma determinada reação, a proporção de dXTP: ddXTP (onde X pode ser qualquer uma das quatro bases) é mantida em aproximadamente 100:1, de modo que a probabilidade de término em um determinado X na cadeia nascente é de cerca de 1/100. Isto produz uma população de fragmentos que terminam em todos os potenciais sítios de término (X) dentro de uma distância de algumas centenas de nucleotídeos do *primer* original de término.

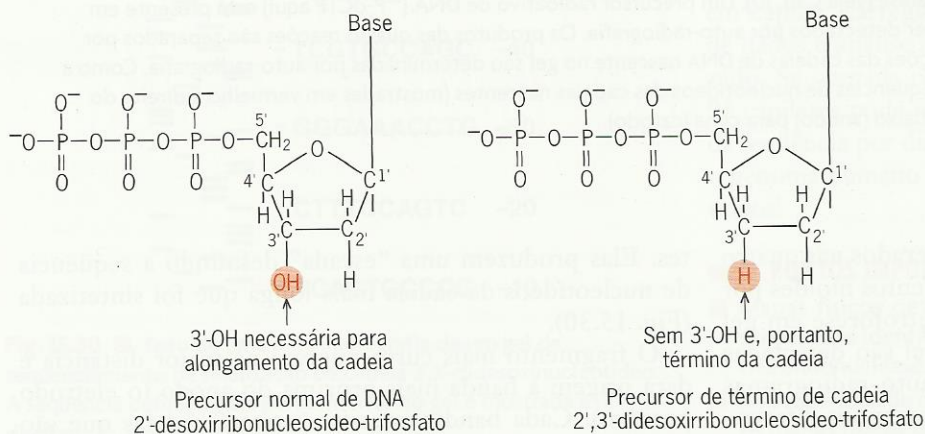


Fig. 15.28 ■ Comparação das estruturas do precursor de DNA normal 2'-desoxirribonucleosídeo-trifosfato e o finalizador de cadeia 2',3'-didesoxirribonucleosídeo-trifosfato usado em reações de sequenciamento de DNA.

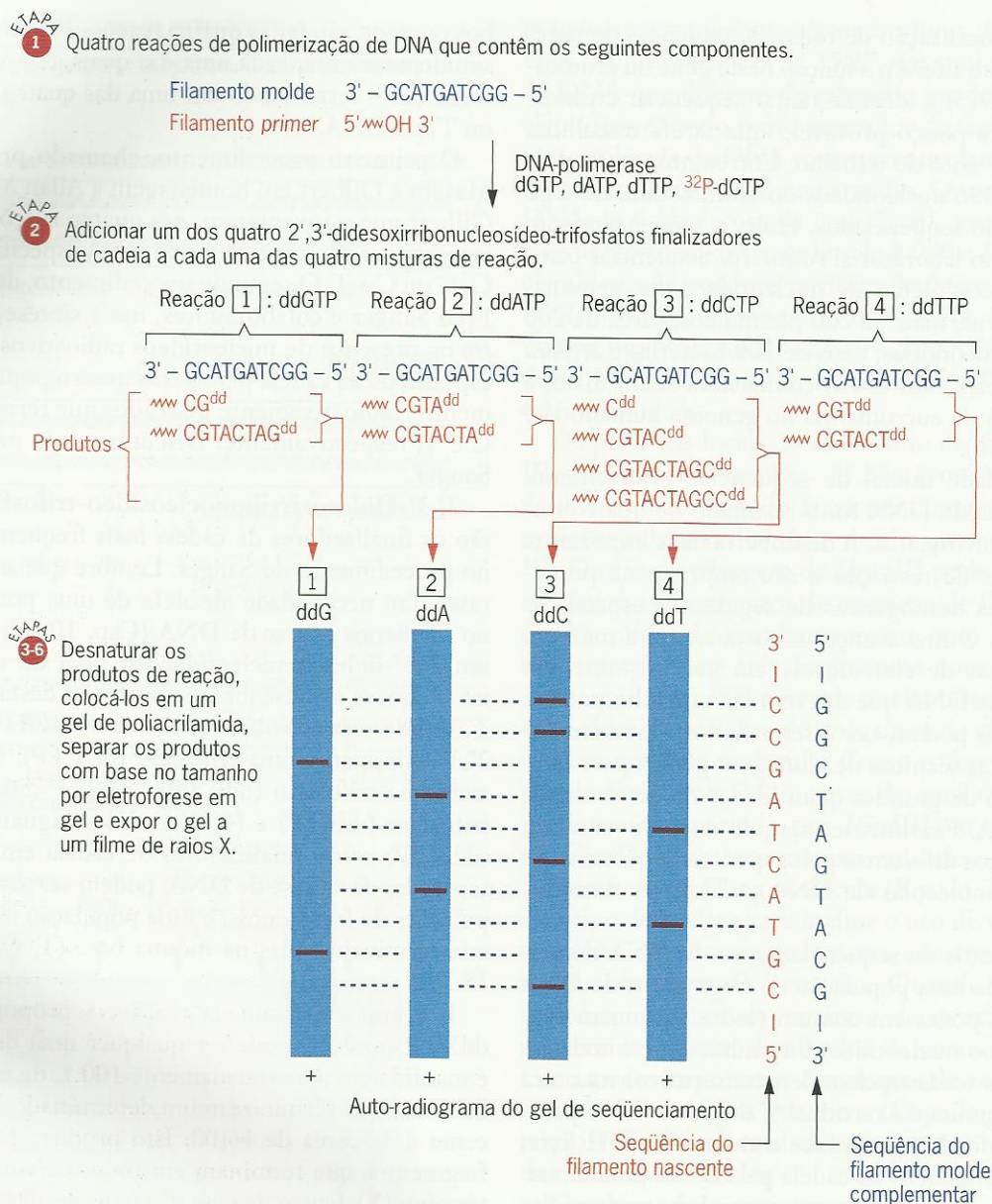


Fig. 15.29 ■ Seqüenciamento de DNA pelo procedimento de término de cadeia com 2',3'-didesoxinucleosídeo-trifosfato. São feitas quatro reações em paralelo, cada uma das quais contém um dos quatro finalizadores de cadeia 2',3'-dideoxi: ddGTP, ddATP, ddCTP e ddTTP. Todas as quatro misturas de reação contém os componentes necessários para síntese de DNA *in vitro*, incluindo um filamento *primer* que se helicoidiza ao filamento molde. Este filamento *primer* determina a ponta 5' comum de todos os produtos. Ele tem uma 3'-OH livre e se estende no sentido 5' → 3' pela DNA-polimerase (veja Cap. 10). Um precursor radioativo de DNA (³²P-dCTP aqui) está presente em cada reação de modo que os produtos podem ser detectados por auto-radiografia. Os produtos das quatro reações são separados por eletroforese em gel de poliácridamida, e as posições das cadeias de DNA nascente no gel são determinadas por auto-radiografia. Como a cadeia mais curta migra na maior distância, as seqüências de nucleotídeos das cadeias nascentes (mostradas em vermelho à direita do auto-radiograma) são obtidas lendo-se o gel de baixo (anodo) para cima (catodo).

Depois que os fragmentos de DNA gerados nas quatro reações paralelas são liberados dos filamentos moldes por desnaturação, eles são separados por eletroforese em gel de poliácridamida, e suas posições no gel são detectadas por auto-radiografia. As bandas nos auto-radiogramas correspondem a cadeias radioativas de tamanhos diferen-

tes. Elas produzem uma “escada” definindo a seqüência de nucleotídeos da cadeia mais longa que foi sintetizada (Fig. 15.30).

O fragmento mais curto migrará na maior distância e dará origem à banda mais próxima do anodo (o eletrodo positivo). Cada banda sucessiva conterá cadeias que são

um nucleotídeo maior do que as cadeias na banda precedente da escada. O nucleotídeo 3'-terminal da cadeia em cada banda será o didesoxinucleotídeo finalizador de cadeia presente na mistura de reação (1, 2, 3 ou 4) no qual esta cadeia específica foi produzida (Fig. 15.29). Lendo a escada produzida por auto-radiografia dos géis de poliacrilamida usados para separar os fragmentos gerados em cada uma das quatro reações paralelas, a sequência completa

de nucleotídeos de uma cadeia de DNA pode ser determinada. Isto é ilustrado na Fig. 15.29 para uma sequência hipotética de nucleotídeos. Um auto-radiograma real de um gel de seqüenciamento de término de cadeia de dideoxinucleotídeos é mostrado na Fig. 15.30. Em condições ótimas, longas seqüências de várias centenas de nucleotídeos podem ser determinadas a partir de um único gel de seqüenciamento.

Hoje, todo o seqüenciamento de DNA em larga escala é feito com máquinas automatizadas de seqüenciamento de DNA que usam o procedimento dideoxi de término de cadeia recém-descrito, mas com modificações. Máquinas diferentes de seqüenciamento de DNA usam protocolos ligeiramente diferentes. Entretanto, as principais diferenças entre o procedimento de seqüenciamento em gel ilustrado na Fig. 15.29 e o seqüenciamento automatizado de DNA são (1) o uso de corantes fluorescentes, em vez de isótopos radioativos, para detectar as cadeias de DNA; (2) a separação dos produtos de todas as quatro reações dideoxi de término de cadeia por eletroforese através de um único gel ou tubo capilar; (3) o uso de fotocélulas para detectar a fluorescência dos corantes à medida que passam pelo gel ou tubo capilar; e (4) a transferência direta do que passou pela fotocélula para um computador, que automaticamente analisa, registra e imprime os resultados.

Um corante fluorescente diferente é usado para marcar os produtos de cada uma das quatro reações dideoxi de seqüenciamento de término de cadeia. Como resultado, os produtos das quatro reações podem ser distinguidos por sua fluorescência à medida que passam através do gel ou tubo capilar. Os corantes fluorescentes podem ser acoplados aos *primers* usados nas reações de seqüenciamento ou acoplados diretamente aos trifosfatos de dideoxinucleosídeos finalizadores de cadeia. A Fig. 15.31 compara o procedimento de seqüenciamento de DNA automatizado com o procedimento de gel descrito na Fig. 15.29 e mostra uma impressão de computador dos resultados de seqüenciamento automatizado de um curto segmento de DNA.

As máquinas de seqüenciamento de DNA automatizadas podem fazer 96 separações simultâneas de eletroforese em capilar, carregamento de amostra, eletroforese, coleta de dados e análise de dados de modo totalmente automatizado. Se operada continuamente, uma única máquina com 96 capilares pode seqüenciar mais de 100.000 nucleotídeos da sequência por dia. Embora isto pareça muito, lembre que o genoma humano contém três bilhões de pares de nucleotídeos!



Fig. 15.30 ■ Foto de uma auto-radiografia de um gel de seqüenciamento de finalização de cadeia 2',3'-dideoxinucleotídeo. A sequência definida pela parte inferior do gel é mostrada à direita.

PONTOS IMPORTANTES

- Mapas físicos detalhados de moléculas de DNA podem ser preparados identificando-se os sítios que são clivados por várias endonucleases de restrição.
- As seqüências de nucleotídeos de moléculas de DNA fornecem os mapas físicos de genes e cromossomos.

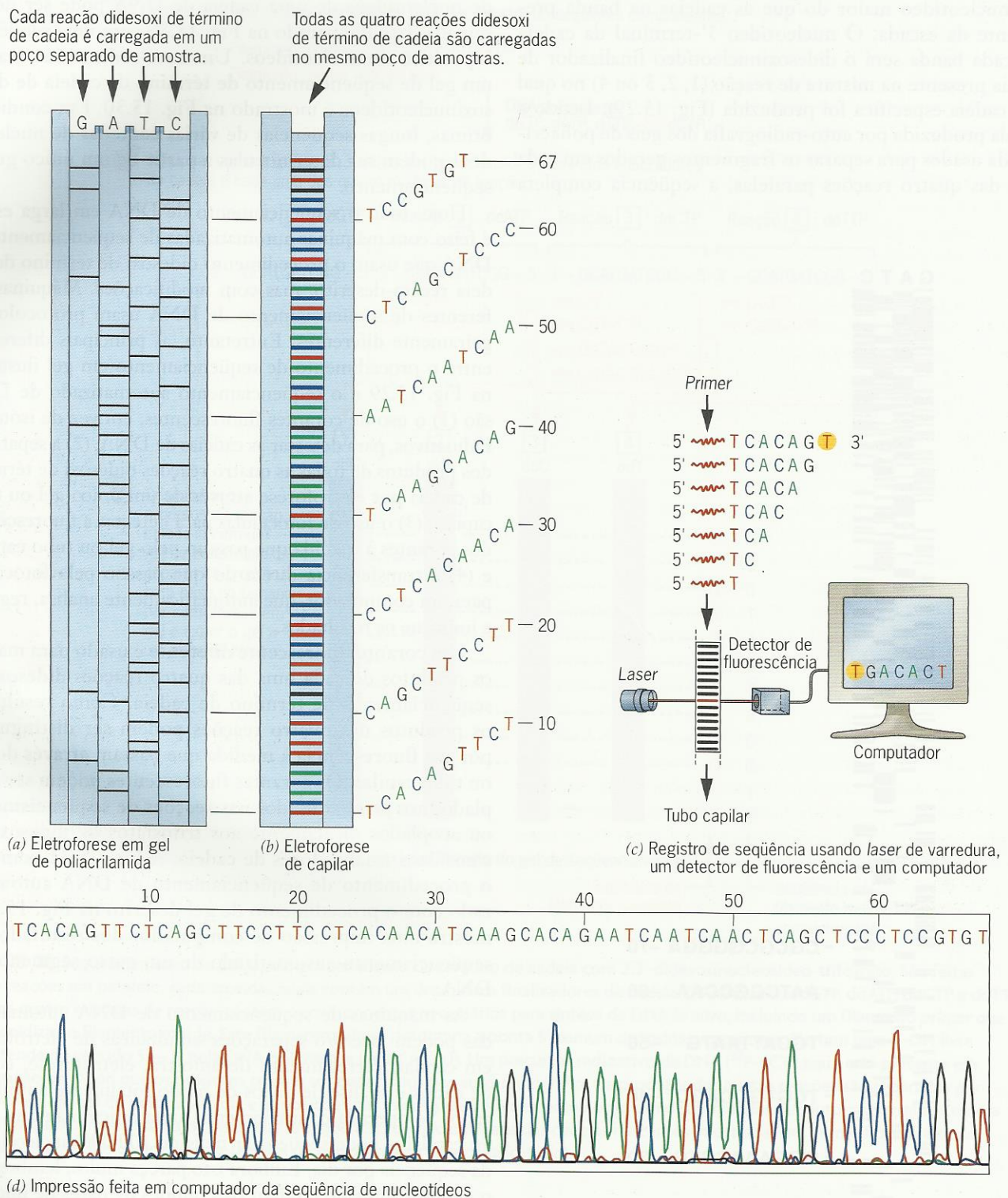
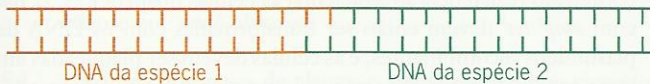


Fig. 15.31 ■ Comparação do gel padrão (a) e procedimentos de seqüenciamento automatizado de DNA em gel de capilaridade (b). (c) No procedimento de seqüenciamento automatizado, os produtos das quatro reações didesoxi de término de cadeia são marcados com quatro corantes que fluorescem em diferentes comprimentos de onda. Os produtos de todas as quatro reações são desnaturados, carregados no mesmo poço e separados com base no tamanho por eletroforese de capilaridade em gel. A fluorescência de cada corante é detectada por uma fotocélula (detector de fluorescência) à medida que cadeias nascentes de DNA se movem através de um feixe de laser e são registradas por um computador. (d) Uma impressão por computador dos resultados de um seqüenciamento automatizado de DNA, mostrando a seqüência de nucleotídeos de um curto segmento de DNA.

Exercícios Básicos Ilustram a análise genética básica.

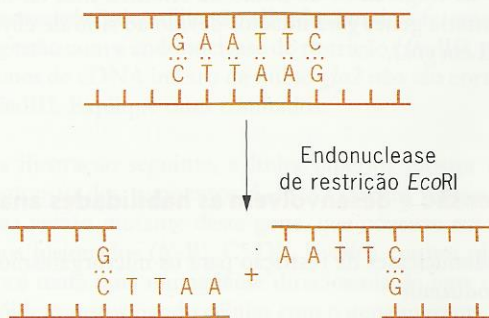
1. O que é uma molécula de DNA recombinante?

Resposta: Uma molécula de DNA recombinante é construída *in vitro* a partir de partes de duas moléculas diferentes de DNA, em geral moléculas de duas espécies diferentes.



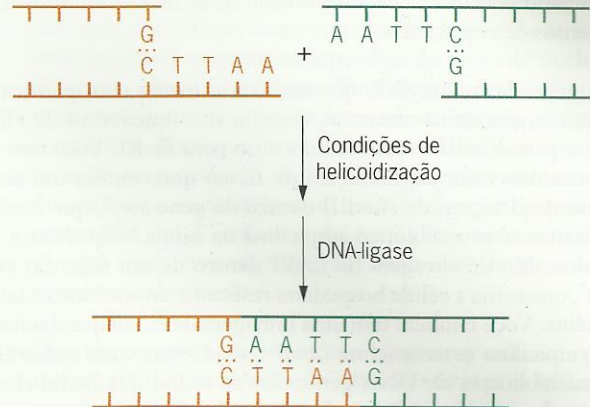
2. O que são endonucleases de restrição?

Resposta: Endonucleases de restrição são enzimas que cortam moléculas de DNA de modo específico para sequência de modo que todos os fragmentos produzidos têm as mesmas seqüências de nucleotídeos em suas pontas. Muitas enzimas de restrição fazem cortes desenhados em seqüências palindrômicas de DNA, produzindo fragmentos com pontas unifilamentares complementares, como mostrado adiante.



3. Como as endonucleases de restrição são usadas para construir moléculas de DNA recombinante *in vitro*?

Resposta: Se moléculas de DNA de fontes diferentes (talvez espécies diferentes) são ambas digeridas com uma endonuclease de restrição que reconhece uma seqüência palindrômica de DNA e faz cortes desenhados nos dois filamentos, os fragmentos resultantes terão pontas unifilamentares complementares. Se esses fragmentos são misturados, as pontas complementares irão parear, e a adição de DNA-ligase produzirá moléculas de DNA recombinante, como mostrado adiante.



4. Por que a reação em cadeia da polimerase (PCR) é uma ferramenta muito poderosa para uso em análises de DNA?

Resposta: Como a PCR amplifica exponencialmente seqüências de DNA, podem ser obtidas grandes quantidades de seqüências específicas começando com apenas uma ou algumas moléculas. Se o início é com uma única molécula de DNA, 10 ciclos de replicação produzirão 1.024 duplas hélices de DNA e 20 ciclos produzirão 1.048.576.

5. Como são usados os 2',3'-didesoxirribonucleosídeo-trifosatos nos protocolos de seqüenciamento de DNA?

Resposta: Os 2',3'-didesoxirribonucleosídeo-trifosatos funcionam como finalizadores específicos de síntese de DNA. Quando um 2',3'-didesoxirribonucleosídeo-monofosfato é adicionado à ponta de uma cadeia nascente de DNA, esta cadeia não pode mais ser aumentada pela DNA-polimerase devido à ausência de 3'-OH necessária para ampliação da cadeia. Usando-se as proporções apropriadas de 2'-didesoxirribonucleosídeo-trifosatos em relação a 2',3'-didesoxirribonucleosídeo-trifosatos em reações de síntese de DNA *in vitro*, cadeias de DNA são produzidas terminando em todas as posições de nucleotídeos possíveis. A separação dessas cadeias de DNA nascente por eletroforese em gel e a detecção de suas posições no gel com nucleotídeos radioativos ou corantes fluorescentes são usadas para determinar suas seqüências de nucleotídeos (veja Figs. 15.29 a 15.31).

Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. O genoma humano (haplóide) contém cerca de 3×10^9 pares de nucleotídeos de DNA. Se você faz uma digestão de uma preparação de DNA humano com *NotI*, uma endonuclease de restrição que reconhece e cliva a seqüência octamérica 5'-GCGGCCGC-3', quantos fragmentos diferentes de restrição você esperaria que fossem produzidos? Suponha que as quatro bases (G, C, A e T) são igualmente prevalentes e aleatoriamente distribuídas no genoma humano.

Resposta: Supondo que as quatro bases estão presentes em quantidades iguais e aleatoriamente distribuídas, a chance de

um determinado nucleotídeo específico ocorrer em um determinado sítio é $\frac{1}{4}$. A chance de uma determinada seqüência específica de dinucleotídeo (p. ex., AG) ocorrer é $\frac{1}{4} \times \frac{1}{4} = (\frac{1}{4})^2$, e a probabilidade de uma seqüência específica de octanucleotídeo é $(\frac{1}{4})^8$ ou $1/65.536$. Portanto, *NotI* cortará tais moléculas de DNA em uma média de uma vez em cada 65.536 pares de nucleotídeos. Se uma molécula de DNA é cortada em n sítios, resultarão $n + 1$ fragmentos. Um genoma de 3×10^9 pares de nucleotídeos deve conter cerca de 45.776 ($3 \times 10^9 / 65.536$) sítios de clivagem de *NotI*. Se todo o genoma humano consistir em uma única

molécula de DNA, *NotI* o cortará em $45.776 + 1$ fragmentos. Considerando que estes sítios de clivagem são distribuídos em 24 cromossomos diferentes, a digestão completa do genoma humano com *NotI* deverá produzir cerca de $45.776 + 24$ fragmentos de restrição.

2. O gene de milho *gln2*, que codifica a forma cloroplástica da enzima glutamina-sintetase, contém um único sítio de clivagem para *HindIII*, mas nenhum sítio para *EcoRI*. Você tem um plasmídeo vetor de clonagem de *E. coli* que contém um único sítio de clivagem de *HindIII* dentro do gene *amp^r*, que confere resistência ao antibiótico ampicilina na célula hospedeira e um único sítio de clivagem de *EcoRI* dentro de um segundo gene *tet^r*, que torna a célula hospedeira resistente ao antibiótico tetraciclina. Você também tem uma linhagem de *E. coli* que é sensível a ampicilina e tetraciclina (*amp^s tet^s*). Como você construiria uma biblioteca de DNA genômico de milho que inclui clones levando um gene *gln2* completo?

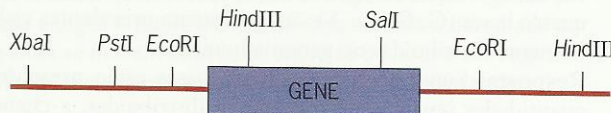
Resposta: O DNA genômico de milho deve ser purificado e digerido com *EcoRI*. Um DNA vetor deve ser similarmente purificado e digerido com *EcoRI*. Os fragmentos de restrição de milho com *EcoRI* e as moléculas de DNA de plasmídeo cortadas com *EcoRI* terão agora pontas unifilamentares complementares (5'-AATT-3'). Os fragmentos de restrição de milho devem em

seguida ser misturados com moléculas de plasmídeo cortadas com *EcoRI* e inseridas covalentemente em moléculas vetorais linearizadas em uma reação dependente de ATP catalisada pela DNA-ligase. A reação de ligação produzirá plasmídeos recombinantes circulares, alguns dos quais conterão insertos de fragmentos de *EcoRI* de milho. A inserção de fragmentos de DNA de milho no sítio de *EcoRI* do plasmídeo perturba o gene *tet^r* de modo que os plasmídeos recombinantes resultantes não mais conferirão resistência a tetraciclina às células hospedeiras. *E. coli* com *amp^s tet^s* devem então ser transformadas com os DNA de plasmídeos recombinantes, e as células devem ser plaqueadas em meio contendo ampicilina para selecionar células transformadas contendo plasmídeos. A maioria das células não serão transformadas e, portanto, não crescerão na presença de ampicilina. As células que crescem em meio contendo ampicilina devem ser aglomeradas e congeladas a -80°C em 20% de glicerol. Essa coleção de células que abrigam fragmentos diferentes de *EcoRI* do genoma de milho representa uma biblioteca de clones que deve conter clones com um gene *gln2* intacto, pois este gene não contém sítio de clivagem de *EcoRI*. Note que o sítio de *HindIII* do vetor poderia ser usado para construir uma biblioteca genômica similar de fragmentos de milho de *HindIII*, mas tal biblioteca não conteria genes *gln2* intactos devido ao sítio de clivagem de *HindIII* em *gln2*.

Questões e Problemas

Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 15.1. (a) De que modos a introdução de moléculas de DNA recombinante em células hospedeiras é similar à mutação? (b) De que modos é diferente?
- 15.2. Abaixo são mostrados quatro filamentos únicos diferentes de DNA. Qual deles, em sua forma bifilar, você esperaria que seja cortado por uma endonuclease de restrição?
- ACTCCAGAATTCACCTCCG
 - GCCTCATTCGAAGCCTGA
 - CTCGCCAATTGACTCGTC
 - ACTCCACTCCCGACTCCA
- 15.3. Se a sequência de pares de bases ao longo de uma molécula de DNA ocorre de modo estritamente aleatório, qual a frequência esperada de uma sequência de reconhecimento de uma enzima de restrição específica com o tamanho de (a) quatro e (b) seis pares de bases?
- 15.4. De que modos as endonucleases de restrição diferem de outras endonucleases?
- 15.5. Que valor têm as tecnologias de DNA recombinante e clonagem gênica para os geneticistas?
- 15.6. O que determina os sítios nos quais as moléculas de DNA serão cortadas por uma endonuclease de restrição?
- 15.7. As endonucleases de restrição constituem valiosas ferramentas para biólogos. Entretanto, os genes que codificam enzimas de restrição obviamente não evoluíram para servir de ferramentas para cientistas. Que possível valor têm as endonucleases de restrição para os microrganismos que as produzem?
- 15.8. Por que o DNA de um microrganismo não é degradado por uma endonuclease de restrição que ele produz, muito embora seu DNA contenha sequências de reconhecimento normalmente clivadas pela endonuclease?
- 15.9. Um dos procedimentos para clonagem de segmentos exógenos de DNA tira proveito de endonucleases de restrição como *HindIII* (veja Quadro 15.1) que produz pontas unifilamentares complementares. Tais enzimas produzem pontas complementares idênticas em DNA exógenos clivados e em DNA vetores nos quais são inseridos DNA exógenos. Que vantagem importante esta estratégia de clonagem tem em relação a procedimentos que usam transferase terminal para sintetizar pontas unifilamentares complementares em DNA exógenos e DNA vetores *in vitro*?
- 15.10. Você está trabalhando como parte de uma equipe de pesquisa que estuda a estrutura e a função de um determinado gene. Seu trabalho é clonar o gene. Está disponível um mapa de restrição para a região do cromossomo no qual está situado o gene; o mapa é o seguinte:



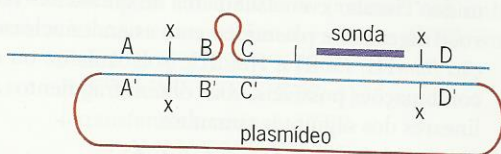
Sua primeira tarefa é preparar uma biblioteca de DNA genômico contendo clones que levam todo o gene. Descreva como você prepararia tal biblioteca no plasmídeo vetor pUC118 (veja a Fig. 15.17), indicando que enzimas de restrição, meio e células hospedeiras você usaria.

15.11. Compare as seqüências de pares de nucleotídeos de clones de DNA genômico e clones de cDNA de genes específicos de plantas e animais superiores. Qual a diferença mais freqüente que você observa?

15.12. A maioria dos genes de plantas e animais que foram clonados logo após o desenvolvimento de tecnologias de DNA recombinante foram genes que codificam produtos que são sintetizados em grandes quantidades em células especializadas. Por exemplo, cerca de 90% das proteínas produzidas em hemácias maduras de mamíferos consistem em cadeias de globina α e β , e os genes de globina estavam entre os primeiros genes clonados de mamíferos. Por que genes deste tipo eram tão prevalentes entre os primeiros genes eucarióticos que foram clonados?

15.13. Clones genômicos do gene de glutamina-sintetase de cloroplastos (*glu2*) de milho são clivados em dois fragmentos por digestão com a endonuclease de restrição *Hind*III, enquanto clones de cDNA inteiro de milho *glu2* não são cortados por *Hind*III. Explique estes resultados.

15.14. Na ilustração seguinte, a linha superior mostra um gene composto dos segmentos A–D. O círculo inferior mostra uma versão mutante deste gene, que consiste em dois trechos fusionados (A'–B', C'–D'), levados em um plasmídeo. Você tenta uma mutagênese direcionada de uma célula diplóide transformando células com o gene mutante clonado. O diagrama seguinte mostra o pareamento desejado do plasmídeo e o cromossomo logo antes de recombinação.



Você prepara o DNA das células, digere-o com uma enzima que corta em x e hibridiza o DNA clivado com a sonda mostrada acima. O diagrama abaixo mostra uma transferência de Southern dos possíveis resultados. (a) Qual coluna mostra fragmentos produzidos pelo DNA na célula antes da transformação? (b) Que coluna mostra os fragmentos produzidos a partir do DNA na célula na qual ocorreu a mutagênese direcionada prevista? (c) Quais destes padrões de transferência devem ser esperados se ocorrerem dois *crossovers*, um entre A e B e o outro entre C e D?

1	2	3	4	5
		—	—	—
—	—	—	—	—
	—		—	—
				—

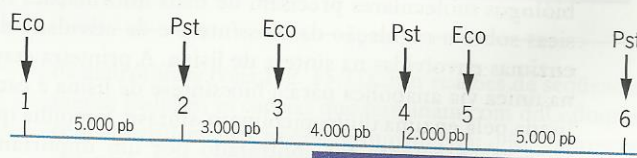
15.15. Você está estudando um gene em *E. coli* que é expresso a 45°C mas não a 37°C. Você mostrou que a regulação deste gene ocorre no nível da transcrição. Além disso, você isolou uma proteína necessária para indução da transcrição do gene em alta temperatura e demonstrou que ela se liga a uma seqüência octamérica específica de pares de nucleotídeos antecedente (5') ao gene. Você clonou o gene completo mais a seqüência reguladora antecedente em um plasmídeo e você tem disponível um mutante de *E. coli* que leva uma deleção do gene todo. Agora você quer determinar quais dos pares de nucleotídeos no sítio de ligação da proteína octamérica em 5' estão envolvidos na interação com (ligação de) a proteína regulatória. Como você determina isto experimentalmente?

15.16. Que vantagem importante tem a reação em cadeia da polimerase (PCR) em relação a outros métodos para análise da estrutura e da função dos ácidos nucleicos?

15.17. Alguns tipos de análises moleculares são facilitados pela disponibilidade de grandes quantidades de DNA unifilar puro a partir de uma molécula de DNA normalmente bifilar. Entretanto a separação e a purificação dos dois filamentos separados de uma determinada dupla hélice geralmente são difíceis de obter pelas técnicas bioquímicas padrões. Como os biólogos moleculares tiram proveito de um mecanismo biológico natural para a purificação de filamentos separados de DNA?

15.18. Quase todos os vetores sofisticados de clonagem – fagos, plasmídeos ou híbridos fago-plasmídeo – hoje em uso têm um componente (além de uma origem de replicação) em comum. Qual é este componente e qual a sua função?

15.19. O desenho abaixo mostra um mapa de restrição de um segmento de uma molécula de DNA. *Eco* refere-se a localizações onde a endonuclease de restrição *Eco*RI corta o DNA, e *Pst* refere-se a locais onde a enzima de restrição *Pst*I corta o DNA. Os sítios potenciais de restrição são numerados de 1 a 6. As distâncias entre os sítios de restrição são mostradas na escala de baixo em pares de bases (pb). A linha grossa representa a parte da molécula que tem homologia com a sonda.



- Suponha que o indivíduo 1 tem sítios de restrição de 1 a 6. Se o DNA for digerido com *Pst*I, quais os tamanhos esperados dos fragmentos de DNA que irão hibridizar com a sonda?
- Suponha que o indivíduo 2 tem uma mutação que elimina o sítio 4. Se o DNA for digerido com *Pst*I, quais os tamanhos esperados dos fragmentos de DNA que irão hibridizar com a sonda?
- Suponha que o indivíduo 3 tem uma mutação que elimina o sítio 5. Se o DNA é digerido com *Pst*I, quais os tamanhos esperados dos fragmentos de DNA que irão hibridizar com a sonda?

- (d) Se o DNA do indivíduo 1 é digerido com *Pst*I e *Eco*RI, quais são os tamanhos esperados dos fragmentos de DNA que irão hibridizar com a sonda?
- (e) Se o DNA do indivíduo 3 é digerido com *Pst*I e *Eco*RI, quais os tamanhos esperados dos fragmentos de DNA que irão hibridizar com a sonda?

15.20. Uma ferramenta importante de biólogos moleculares envolve a transferência de proteínas que foram separadas por eletroforese em gel para membranas de nitrocelulose e a detecção de proteínas específicas nas membranas por meio de anticorpos e marcadores radioativos acoplados ou reações enzimáticas acopladas. Quando é usado este procedimento, a visualização das bandas de proteínas resultantes é chamada de transferência *western*. Qual o significado do nome “transferência *western*”?

15.21. O gene da fibrose cística (*CF*) (localização: cromossomo 7, região q31) foi clonado e seqüenciado, e estudos de pacientes mostraram que cerca de 70% deles são homozigotos para um alelo mutante *CF* que tem uma deleção de três pares de nucleotídeos específicos (equivalente a um códon). Essa deleção resulta na perda de uma fenilalanina na posição 508 no produto do gene *CF*. Suponha que você é um consultor genético responsável por orientar famílias com *CF* em seus heredogramas quanto ao risco de *CF* entre sua prole. Como você pode triar supostos pacientes com *CF* e seus genitores e parentes quanto à presença do gene mutante *CF* Δ F508? O que a detecção deste gene mutante em uma família lhe permite dizer sobre as chances de que *CF* ocorra novamente na família?

15.22. Cereais são importantes fontes de alimento para seres humanos e outros animais em muitas regiões do mundo. Entretanto, a maioria dos cereais contém quantidades inadequadas de determinados aminoácidos que são essenciais para animais monogástricos como os humanos. Por exemplo, o milho contém quantidades insuficientes de lisina, triptofano e treonina. Assim, uma importante meta para geneticistas de planta é produzir variedades de milho com aumento do conteúdo de lisina. Como pré-requisito para a construção de milho rico em lisina, biólogos moleculares precisam de mais informações básicas sobre a regulação da biossíntese e da atividade das enzimas envolvidas na síntese de lisina. A primeira etapa na única via anabólica para a biossíntese de lisina é catalisada pela enzima diidropicolinato-sintase. Suponha que você recentemente foi contratado por um importante instituto de pesquisas vegetais nos EUA que lhe incumbiu de isolar um clone da seqüência de ácido nucleico que codifica a diidropicolinato-sintase em milho. Descreva resumidamente quatro enfoques diferentes que você pode adotar na tentativa de isolar tal clone e incluir pelo menos uma abordagem genética.

15.23. (a) Que procedimento experimental é feito em análises de Southern, *northern* e *western*? (b) Qual a principal diferença entre as análises de Southern, *northern* e *western*?

15.24. Você isolou um clone de cDNA que codifica uma proteína de interesse em um eucarionte superior. Esse clone de

cDNA não é cortado pela endonuclease de restrição *Eco*RI. Quando esse cDNA é usado como uma sonda radioativa para análise de hibridização de DNA genômico digerido com *Eco*RI, são vistas três bandas radioativas na transferência de Southern resultante. Isto indica que o genoma do eucarionte em questão contém três cópias do gene que codifica a proteína de interesse?

15.25. Uma molécula linear de DNA é submetida a digestões única e dupla com endonucleases de restrição, e são obtidos os seguintes resultados:

Enzimas	Tamanhos dos fragmentos (em kb)
<i>Eco</i> RI	2,9, 4,5, 7,4, 8,0
<i>Hind</i> III	3,9, 6,0, 12,9
<i>Eco</i> RI e <i>Hind</i> III	1,0, 2,0, 2,9, 3,5, 6,0, 7,4

Desenhe o mapa de restrição definido por estes dados.

15.26. Uma molécula de DNA é submetida a digestões únicas e duplas com enzimas de restrição, e os produtos são separados por eletroforese em gel. Os resultados são mostrados abaixo (os tamanhos dos fragmentos estão em kb):

<i>Eco</i> RI e			<i>Eco</i> RI e <i>Hind</i> III e		
<i>Eco</i> RI	<i>Hind</i> III	<i>Hind</i> III	<i>Bam</i> HI	<i>Bam</i> HI	<i>Bam</i> HI
8	5	12	6	6	6
4	4		6	4	5
	3			2	1

Desenhe o mapa de restrição desta molécula de DNA.

15.27. Você está estudando uma molécula de DNA de um plasmídeo circular com 10,5 pares de quilobases (kb). Quando você digere este plasmídeo com as endonucleases de restrição *Bam*HI, *Eco*RI e *Hind*III, isoladamente ou em todas as combinações possíveis, você obtém fragmentos de restrição lineares dos seguintes tamanhos:

Enzimas	Tamanhos dos fragmentos (em kb)
<i>Bam</i> HI	7,3, 3,2
<i>Eco</i> RI	10,5
<i>Hind</i> III	5,1, 3,4, 2,0
<i>Bam</i> HI + <i>Eco</i> RI	6,7, 3,2, 0,6
<i>Bam</i> HI + <i>Hind</i> III	4,6, 2,7, 2,0, 0,7, 0,5
<i>Eco</i> RI + <i>Hind</i> III	4,0, 3,4, 2,0, 1,1
<i>Bam</i> HI + <i>Eco</i> RI + <i>Hind</i> III	4,0, 2,7, 2,0, 0,7, 0,6, 0,5

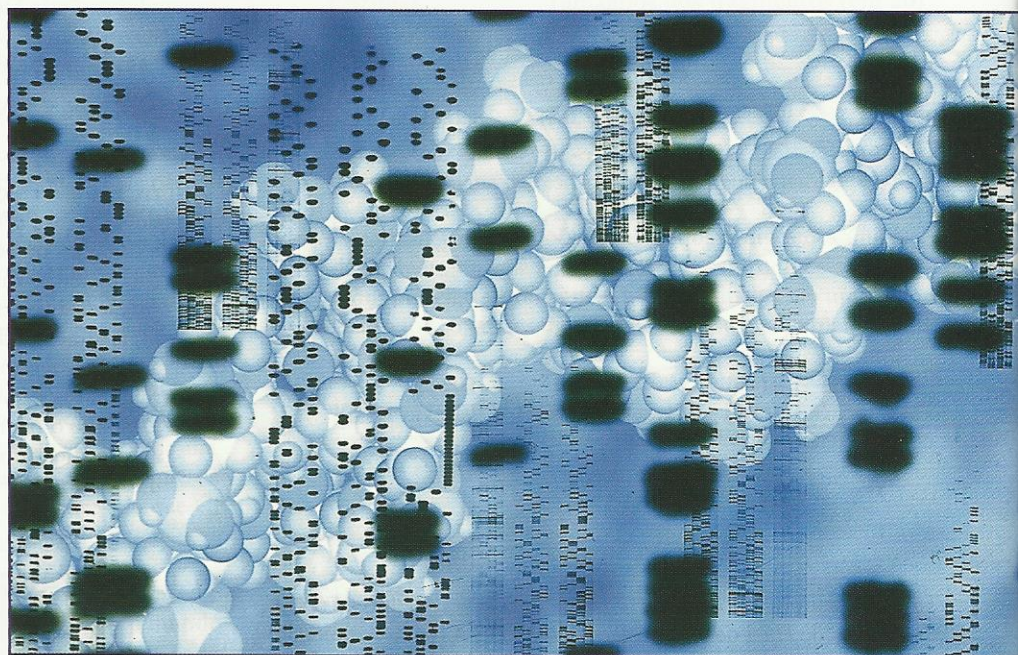
Desenhe um mapa de restrição para o plasmídeo que se ajusta a seus dados.

15.28. O seguinte é um segmento de um auto-radiograma de um gel de seqüenciamento didesoxi. A origem é em cima. Qual a seqüência (incluindo a polaridade) do DNA unifilar que serve como um molde para gerar este padrão? (Use todas as bandas no gel.)

Genômica

Tópicos

- Genômica: Uma Visão Geral
- Correlação dos Mapas Genético, Citológico e Físico de Cromossomos
- Clonagem de Genes Baseada em Posição de Mapa
- O Projeto do Genoma Humano
- Análises de RNA e Proteínas do Funcionamento do Genoma
- Genômica Comparativa



Auto-radiogramas, vistos de perto e em tamanho total, mostrando as seqüências de nucleotídeos de DNA superpostas sobre um modelo preenchido do DNA (mostrado em branco).

Levantamento de Genes Humanos na Islândia

Uma importante pesquisa de genes humanos com significativo valor farmacológico está ocorrendo em um local um tanto inesperado, a remota ilha da Islândia. Devido a sua história e seu isolamento geográfico, as 270.000 pessoas da Islândia constituem uma fonte singular para estudos genéticos. Eles são os descendentes bem homogêneos dos Vikings que se instalaram nessa ilha há mais de 1.100 anos. Tal homogeneidade foi realçada por dois “gargalos genéticos” durante os quais a população da Islândia foi acentuadamente reduzida. Durante o século quinze, a população caiu de cerca de 70.000 para cerca de 25.000 quando a peste bubônica devastou a ilha. Durante os anos 1700, a população caiu abaixo dos 50.000 em três ocasiões devido a fome e doenças causadas em parte pela erupção do vulcão Hekla. Assim, o *pool* de genes humanos da Islândia é muito mais homogêneo que os *pools* gênicos da maioria de outras populações. Além disso, o serviço de saúde nacional da Islândia tem mantido ótimos registros médicos familiares desde 1915.

Em 1997, Kari Stefansson, geneticista de Harvard, reconheceu a singularidade do *pool* gênico humano da Islândia e seus registros médicos. Ele voltou a sua terra natal para desenvolver uma empresa privada, a deCODE Genetics, com a meta de identificar genes humanos que levariam ao desenvolvimento de novas drogas farmacêuticas e testes diagnósticos. Em 1998, o governo da Islândia deu à deCODE uma permissão exclusiva para construir e analisar um banco de dados de genética a partir dos registros de saúde do país. O primeiro sucesso da empresa foi a identificação do *tremor essencial familiar*, um gene associado ao tremor nos idosos. Cientistas da deCODE também identificaram genes envolvidos na suscetibilidade a doença cardíaca, acidente vascular cerebral, hipertensão, doença de Alzheimer e câncer de próstata.

Com base nestes resultados, a deCODE Genetics negociou um contrato de US\$200 milhões com a gigante farmacêutica suíça Hoffmann-LaRoche, que deu à empresa suíça direitos exclusivos para qualquer droga ou produtos diagnósticos resultantes do trabalho dos cientistas da deCODE. Para as pessoas da Islândia, o contrato especifica que a Hoffmann-LaRoche deve fornecer sem nenhuma despesa todas as drogas, testes diagnósticos e outros produtos resultantes desta pesquisa. Portanto, pelo menos neste caso, as pessoas que forneceram os dados genéticos e as amostras de DNA para análise serão pessoalmente beneficiadas pelos resultados das pesquisas de uma empresa privada.

A saga da deCODE, entretanto, não progrediu sem controvérsia. O ponto chave nos contínuos debates na Islândia é a questão sobre *consentimento presumido* versus *consentimento informado*. De acordo com a legislação aprovada pelo parlamento da Islândia, os registros de saúde de uma pessoa são automaticamente incluídos no banco de dados a menos que ela ou ele solicite especificamente que os registros sejam excluídos. Esta lei estabelece a prática do “consentimento presumido”. A prática geral no campo da medicina é do “consentimento informado”, em que todos os registros de uma pessoa são considerados privados a menos que a pessoa assine especificamente um documento de consentimento informado dizendo que eles podem ser liberados. Muitos cidadãos e médicos da Islândia fazem objeção à lei de consentimento presumido que regulamenta a deCODE, dizendo que é uma invasão de sua privacidade. Alguns médicos estão recusando submeter os registros de saúde de seus pacientes à deCODE. Entretanto, a lei da Islândia que estabelece a prática do consentimento presumido é muito clara, e o banco de dados da deCODE continua a se expandir, da mesma forma que o debate sobre os procedimentos de coleta e uso de dados genéticos de populações humanas.

Gregor Mendel estudou os efeitos de sete genes sobre características em ervilhas, mas ele não estudou mais que três genes em nenhum cruzamento. Geneticistas de hoje podem estudar a expressão de todos os genes, o genoma inteiro, de um organismo em um único experimento. Em janeiro de 2005, as seqüências completas, ou quase completas, de nucleotídeos dos genomas de mais de 1.000 vírus, vírus defectivos e viróides, 671 plasmídeos, 652 mitocôndrias, 41 cloroplastos, 20 *archaea*, 190 bactérias e 28 eucariontes haviam sido determinadas. A lista de eucariontes cujos genomas haviam sido seqüenciados inclui importantes organismos modelos em genética: a levedura do pão *Saccharomyces cerevisiae*, a mosca-das-frutas *Drosophila melanogaster* e a planta *Arabidopsis thaliana*. Também inclui o protozoário *Plasmodium falciparum*, que causa a forma mais perigosa de malária, e

o mosquito *Anopheles gambiae*, que é o organismo hospedeiro mais responsável pela disseminação desta doença. O bicho-da-seda (*Bombyx mori*), um inseto economicamente importante, está na lista, bem como vários vertebrados: o camundongo (*Mus musculus*), o rato da Noruega (*Rattus norvegicus*), o *Red Jungle Fowl*, um ancestral dos galos domésticos (*Gallus gallus*), o baiacu (*Fugu rubripes*) e nossa própria espécie (*Homo sapiens*).

Uma das metas originais do Projeto do Genoma Humano foi determinar a seqüência completa de nucleotídeos do genoma humano no ano de 2005. O que ocorreu é que dois rascunhos da seqüência – um por um consórcio público e outro por uma empresa privada – foram publicados em fevereiro de 2001 (veja Um Marco na Genética: Dois Rascunhos da Seqüência do Genoma Humano, no final deste capítulo). Uma seqüência quase completa do genoma humano com

99% do DNA eucromático foi liberada em outubro de 2004, um ano inteiro antes da meta original.

As melhorias na tecnologia de seqüenciamento do DNA que ocorreram durante as últimas duas décadas do século vinte permitiram que pesquisadores coletassem grandes quantidades de dados de seqüência. Entretanto, o seqüenciamento nem sempre foi tão fácil. Custou a Robert Holley, ganhador do Prêmio Nobel de 1968, vários anos para determinar a seqüência de 77 nucleotídeos do tRNA de alanina

de levedura. Alguns dos principais avanços na tecnologia do seqüenciamento, bem como alguns dos marcos no estudo dos genomas, são destacados na Fig. 16.1.

Na era atual, grandes quantidades de dados de seqüências acumulam-se diariamente. A maioria destes dados resulta de projetos de pesquisa financiados por agências governamentais – o National Institutes of Health (NIH), o National Science Foundation (NSF) e o Department of Energy (DOE) nos EUA – e agências comparáveis em outros países.



ENFOQUE TÉCNICO: GenBank

Em 1979, Walter Goad, um físico que trabalhava no Los Alamos National Laboratory (LANL) no Novo México, teve a idéia de um banco de dados que conteria todas as seqüências de DNA disponíveis. De 1982 até 1992, Goad e seus colaboradores incorporaram seqüências ao banco de dados, hoje chamado GenBank, e o mantiveram no LANL. Hoje, esse banco de dados é mantido pelo National Center for Biotechnology Information (NCBI), que é parte da National Library of Medicine (NLM) no National Institutes of Health (NIH) em Bethesda, Maryland. O conteúdo do banco de dados cresceu muito desde que Goad e seus colaboradores o criaram (Fig. 1). No final de 1982, o GenBank continha 680.338 pares de nucleotídeos de seqüências de DNA, mas ao final de 2004 ele continha mais de 44 bilhões de pares de nucleotídeos.

Também foram estabelecidos bancos de dados comparáveis ao GenBank na Europa e no Japão. O European Molecular Biology Laboratory (EMBL) Data Library foi criado na Alemanha em 1980, e o DNA DataBank of Japan (DDBJ) foi estabelecido em 1984. GenBank, EMBL e DDBJ subsequentemente se uniram e formaram o International Nucleotide Sequence Database Collaboration, que permite a pesquisadores analisar todos os três bancos de dados simultaneamente.

O desenvolvimento de programas de pesquisa e recuperação de triagem de bancos de dados quanto a seqüências similares às que são colocadas deu aos cientistas uma grande ferramenta de pesquisa. Em particular, o sistema de pesquisa do ENTREZ do NCBI é muito valioso. Este sistema foi distribuído em CD-ROM em 1992, depois como versão da rede em 1993 e finalmente está dis-

ponível, grátis, na Internet desde 1994 (URL: <http://www.ncbi.nlm.nih.gov/entrez>). A quantidade de informações pesquisáveis e que se pode obter disponíveis no *web site* Entrez tem aumentado a cada ano. Ela não só engloba os bancos de dados de seqüências de proteínas como também inclui um enorme banco de dados bibliográficos chamado PubMed que cobre a maioria dos periódicos de medicina e biologia. Hoje, você pode pesquisar todos estes

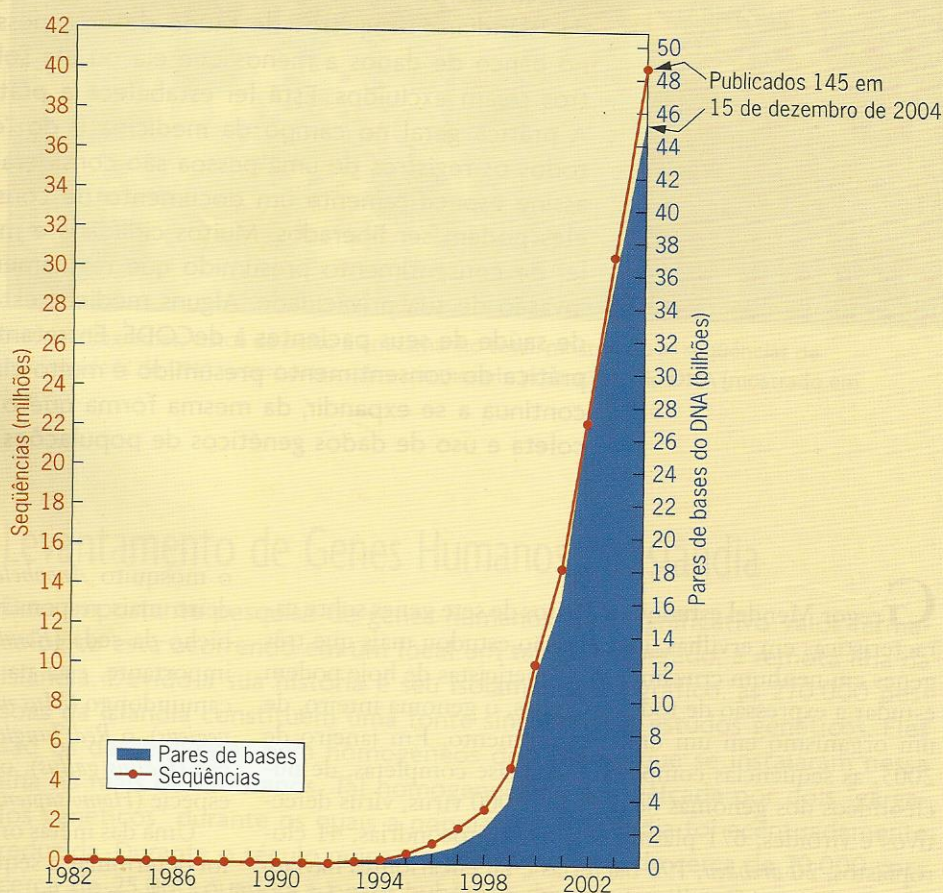


Fig. 1 ■ Crescimento do GenBank desde sua origem em 1982 até 15 de dezembro de 2004. As ordenadas à esquerda e à direita mostram o tamanho da coleção em número de seqüências de DNA (vermelho) e número de pares de nucleotídeos (azul), respectivamente. O número de seqüências diferentes cresceu de 606 no final de 1982 para 40,6 milhões no final de 2004.

Assim, esses dados são informações públicas e estão disponíveis para qualquer um que queira usá-los. As seqüências tornaram-se públicas por meio do estabelecimento de bancos de dados de seqüências que estão disponíveis grátis na web em <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi> (veja Enfoque Técnico: GenBank).

Logicamente, apenas tornar disponíveis os bancos de dados não é suficiente. Devemos ser capazes de extrair informações deles, isto é, “garimpar” os bancos de dados, e

então analisar eficientemente e precisamente a informação extraída. Este processo requer programas de computação que podem procurar as grandes seqüências de DNA nos genomas de interesse. A necessidade de tal programa englobou uma nova disciplina científica chamada bioinformática. Matemáticos, cientistas de computação e biólogos moleculares que trabalham nesta disciplina desenvolveram algoritmos de pesquisa em computador que podem extrair informações do DNA e dados de seqüências de proteínas.



ENFOQUE TÉCNICO: GenBank (continuação)

bancos de dados simultaneamente usando o cruzamento global do banco de dados do NCBI, e a página de procura fornece o número de itens encontrados (isto é, as “entradas”) em cada banco de dados.

Uma discussão de todos os bancos de dados que podem ser pesquisados com o Entrez está além do escopo deste texto. Você é encorajado a visitar o site e explorar alguns de seus bancos de dados. Eles incluem os bancos de dados do PubMed e DNA mencionados previamente e bancos de dados de seqüências de proteínas, estruturas macromoleculares tridimensionais, cromossomos e genes causadores de câncer, seqüências expressas, polimorfismos de um só nucleotídeo, seqüências genômicas inteiras e muito mais.

Vamos fazer uma pesquisa no Entrez para ilustrar como ele funciona. Suponha que você quer determinar a seqüência de nucleotídeos de um segmento de DNA de um organismo de interesse e quer saber se esse DNA já foi seqüenciado ou se é similar a seqüências em qualquer um dos atuais bancos de dados. Um dos modos mais rápidos de obter esta informação é fazer uma pesquisa no BLAST (Basic Local Alignment Search Tool) com sua seqüência de entrada, ou *query*. Começamos na *home page* do NCBI: <http://www.ncbi.nlm.nih.gov/>. Primeiro, selecione “BLAST” na barra de ferramentas. Então, em “Nucleotide” em cima à esquerda, clique em “Quickly search for highly similar sequences (megablast)”, e uma janela similar à mostrada na Fig. 2a irá aparecer.

Suponha que sua seqüência de nucleotídeos é a seguinte:

```
5'-ATGAGAGAAATTCTTCATATTCAAGGAGGTCAGTGCGGAA
ACCAGATCGGAGCTAAGTTCTGGGAAGTTATTGCGGCGAG
CACGGTATTGATCAAACCG-3'
```

Digite ou transfira a seqüência indagada para o box no topo da página do BLAST. Antes de clicar o botão “BLAST!”, role para “Format” e faça algumas mudanças para diminuir a quantidade de informações que você irá receber. Mude “alignment” do “HTML” para “Plain text” e reduza tanto o número de “descrições” quanto o número de “alignments” para 10. Finalmente, mude o “alignment view” para “Pairwise.” Agora, clique o botão “BLAST!”. Seus resultados deverão aparecer em menos de um minuto. Eles devem incluir uma lista de 10 “Sequences producing significant alignments” (descrições de seqüências 5 e 10 são

mostradas na Fig. 2b) e o alinhamento de cada seqüência com sua seqüência indagada (os alinhamentos são mostrados para os números 5 e 10 na Fig. 2c).

As 10 seqüências mais similares nos bancos de dados incluem apenas duas seqüências distintas. As primeiras seis seqüências são todas obtidas independentemente do mesmo gene; as quatro seguintes são seqüências independentes de um segundo gene proximoamente relacionado. Note na Fig. 2c que a seqüência indagada tem um ajuste perfeito com a seqüência 5 e que difere da seqüência nas posições de nucleotídeos 10 e 12. As seqüências 5 e 10 são membros de uma família de genes que codificam um conjunto de proteínas proximoamente correlatas com as mesmas funções ou funções muito similares.

Suponha que você quer saber mais sobre as seqüências identificadas em sua pesquisa. Vamos para a *home page* do NCBI: <http://www.ncbi.nlm.nih.gov/entrez/>. Role para a parte inferior da página e clique em “Go” para uma pesquisa em todos os bancos de dados. Na “Search across databases”, digite o número de acesso para a seqüência 5, “M84706”, e clique em “Go”. A pesquisa levará cerca de 1 segundo, e a página de resumo mostrará as correspondências em “PubMed” e uma de cada em “PubMed Central”, “Nucleotide” e “Gene”. Clique primeiro em “Nucleotide”, e você verá que a seqüência é parte do gene de beta-9-tubulina de *Arabidopsis thaliana*. (A β -tubulina é uma das proteínas estruturais das fibras do fuso anafásico.) Em seguida, clique em “M84706”, e a página resultante dará uma descrição do gene, incluindo a seqüência completa de nucleotídeos.

Agora, voltemos à página de resumo da pesquisa cruzada de bancos de dados e cliquemos em “Pub Med” para ver as equivalências na literatura. Serão mostradas duas referências. Clique na segunda: “Snustad *et al.*”, e a página título e o resumo desta publicação aparecerão na janela. Clique em “Free full text article”, e uma versão PDF da publicação poderá ser baixada para seu computador. Esta referência contém os primeiros relatos das seqüências de nove genes de β -tubulina de *Arabidopsis*.

Esta breve exploração do web site Entrez ilustra o poder e a conveniência do programa e os bancos de dados hoje disponíveis para analisar seqüências de DNA. Sem tais ferramentas, geneticistas teriam enormes dificuldades em dar sentido ao vasto número de seqüências de DNA atualmente disponíveis.



ENFOQUE TÉCNICO: GenBank (continuação)



Nucleotide

Protein

Translations

Retrieve results for
an BID

What is MegaBLAST?

Search

ATGAGAGAAATTCTTCATATTCAAGGAGGTCAGTGCGGAAACCAGATCGGAGCTAAGTTCTC

Load query file from disk

Browse...

Set subsequence

From:

To:

Choose database

nr

Return alignment
endpoints only

Now:

BLAST!

or

Reset query

Reset all

(a) the MegaBLAST web site at NCBI.

Sequences producing significant alignments:

Score
(bits)

E
value

No. 5: [gil166909|gb|M84706.1|ATHTUB9B Arabidopsis thaliana beta-9 tubul.](#)

192

8e-47

No. 10: [gil30689176|ref|NM_122291.2| Arabidopsis thaliana tubulin beta-8](#) . .

110

 $6e-22$

(b) Two of the 10 sequences identified by the MegaBLAST search.

Query: atgagagaaattcttcatattcaaggaggtcagtcgcaaaccagatcggagctaagtctctgggaagtattttcggcgagcacggtattgatcaaaccg 100
M84706.1 atgagagaaattcttcatattcaaggaggtcagtcgcaaaccagatcggagctaagtctctgggaagtattttcggcgagcacggtattgatcaaaccg 227

Query: atgagagaaattcttcatattcaaggaggtcagtgcggaaaccagatcgagagctaagtcttggaagtta-tttgcggcgagcacggtattgat 93
NM_122291.2 atgcgagagattcttcacatacaaggtggccaatgcggaaaccagatcgagagctaagtcttggaagt-agtttgcgccgagcacgggatcgat 172

(c) Nucleotide alignments of the query sequence with two sequences identified by the MegaBLAST search.

Fig. 2 ■ (a) Ilustração de uma pesquisa MegaBLAST do banco de dados “Nucleotídeo” do Entrez. A sequência de 100 nucleotídeos: ATGAGAGAAATCTTCATATTCAAGGAGGTCTAGTCGGAAA-CCAGATCGGAGCTAAGTTCTGGGAAGTTATTGCGGCGAGC-ACGGTATTGATCAAACCG foi usada como indagação. Os resultados da pesquisa foram restritos às 10 sequências mais similares nos bancos de dados. As primeiras seis sequências identificadas são todas idênticas à sequência indagada. Elas são sequências independentes do gene de β -9-tubulina (*TUB9*) de *Arabidopsis thaliana*. As quatro sequências seguintes identificadas são muito similares, mas não idênticas à sequência indagada. Elas são sequências independentes do gene *TUB8* de *Arabidopsis*, que codificam um β -8-tubulina muito similar. (b) As duas sequências distintas identificadas pela pesquisa MegaBLAST. Quanto mais alto o valor, mais proximamente relacionadas as sequências. (c) Alinhamentos da sequência indagada com cada uma das sequências identificadas na pesquisa MegaBLAST. Note a identidade perfeita da sequência M84706.1 (*TUB9*) com a sequência indagada, e a próxima similaridade de NM_122291.2 (*TUB8*) com a sequência indagada. *TUB8* e *TUB9* codificam β -tubulinas muito similares.

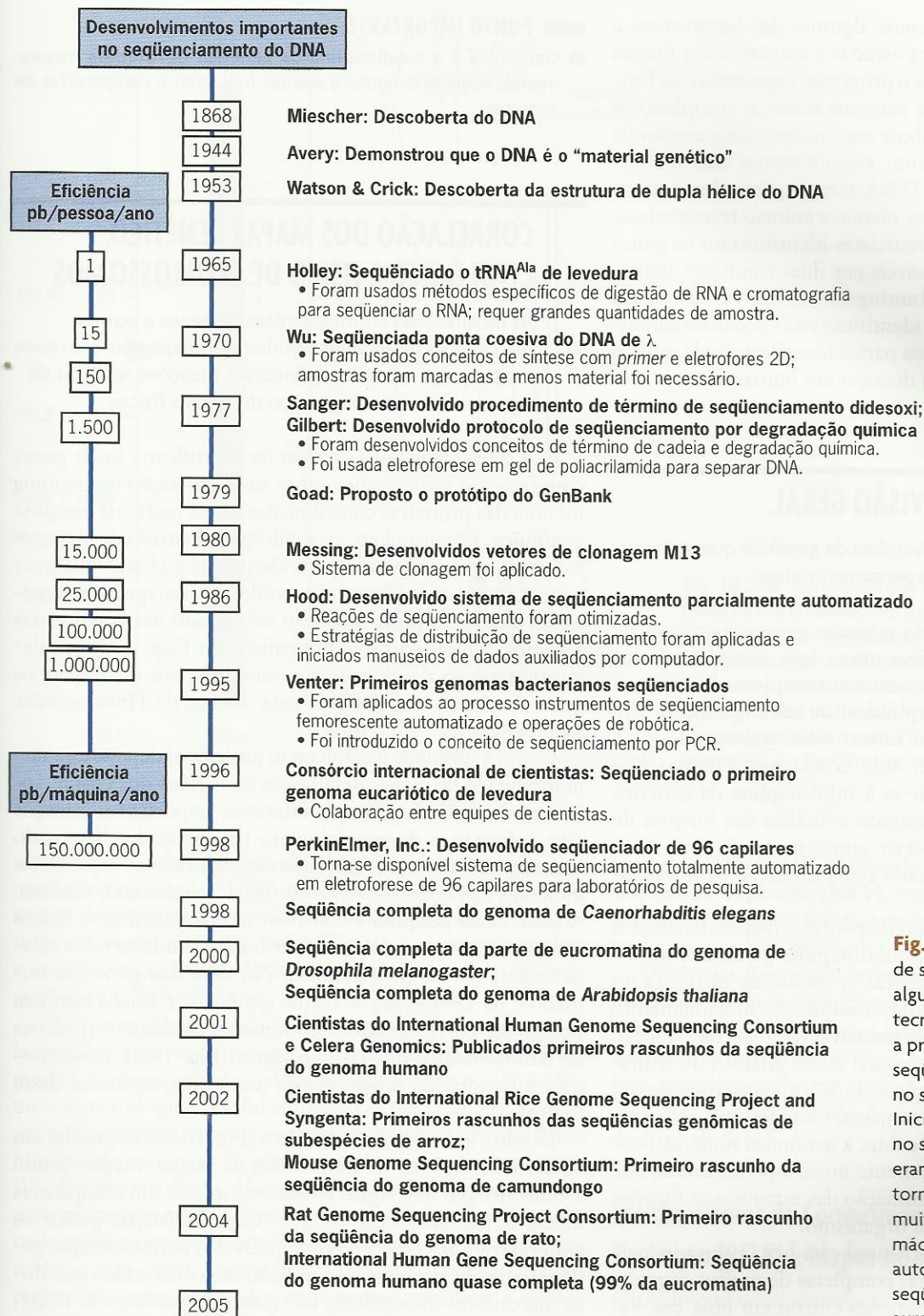


Fig. 16.1 ■ Avanços na eficiência de seqüenciamento do DNA, alguns desenvolvimentos tecnológicos que aumentaram a produtividade de seqüenciadores e alguns marcos no seqüenciamento do DNA. Inicialmente, todas as etapas no seqüenciamento do DNA eram feitas manualmente, tornando-o um processo muito trabalhoso. Entretanto, máquinas de seqüenciamento automatizado hoje substituíram seqüenciamentos humanos, aumentando muito a eficiência.

A disponibilidade de seqüências genômicas inteiras abriu as portas para análises de bioinformática e para estudos funcionais dos genes contidos nessas seqüências. Microarranjos, incluindo os chamados *chips* de genes, permitem que cientistas investiguem a expressão de todos os genes de um organismo simultaneamente (veja Análises de RNA e Prote-

ínas do Funcionamento do Genoma, neste capítulo). Outros procedimentos usam seqüências conhecidas de nucleotídeos para dissecar vias metabólicas por "nocaute" ou desligamento da expressão de genes (veja Genética Reversa, no Cap. 17, e O Camundongo: Mutagênese Inercial, Transgênicos e Mutações Nocaute, no Cap. 22).

Neste capítulo, discutiremos algumas das ferramentas e técnicas que são usadas para estudar a estrutura e a função dos genomas, examinaremos o progresso espetacular do Projeto do Genoma Humano e veremos como as comparações dos genomas podem contribuir para nossa compreensão da evolução. No capítulo seguinte, examinaremos outros avanços técnicos – *fingerprint* de DNA, terapia gênica humana e a produção de microrganismos, plantas e animais transgênicos. Veremos também como geneticistas identificaram os genes defeituosos que são responsáveis por duas condições humanas trágicas, a doença de Huntington e a fibrose cística. Os procedimentos usados para identificar estes genes tornaram-se paradigmas metodológicos para a identificação de muitos outros genes relacionados a doenças em humanos.

GENÔMICA: UMA VISÃO GERAL

Genômica é uma subdisciplina da genética que enfoca a estrutura e a função de genomas inteiros.

Geneticistas têm usado o termo *genoma* por mais de sete décadas para se referir a uma cópia completa da informação genética ou um conjunto completo de cromossomos (monoplóides ou haplóides) de um organismo. Em contraste, **genômica** é um termo relativamente novo. A palavra *genômica* parece ter sido criada por Thomas Roderick em 1986 para referir-se à subdisciplina da genética de mapeamento, seqüenciamento e análise das funções de genomas inteiros e para servir como nome de uma nova publicação, *Genomics*, dedicada à comunicação de novas informações nesta área.

À medida que se tornaram disponíveis mapas detalhados e seqüências de genomas, a subdisciplina genômica foi dividida em **genômica estrutural**, o estudo da estrutura do genoma, **genômica funcional**, o estudo do funcionamento do genoma, e **genômica comparativa**, o estudo da evolução do genoma. A genômica funcional inclui análises do **transcriptoma**, o conjunto completo de RNA transcritos de um genoma, e o **proteoma**, o conjunto completo de proteínas codificadas pelo genoma. De fato, a genômica funcional engloba uma disciplina inteiramente nova, a **proteômica**, que tem como sua meta a determinação das estruturas e funções de todas as proteínas em um organismo.

Enquanto a genômica estrutural está bem avançada, pela disponibilidade de seqüências completas de muitos organismos, a genômica funcional apenas entrou em uma fase explosiva de crescimento. As novas tecnologias de hibridização e *chips* gênicos permitem que pesquisadores monitorem a expressão de genomas inteiros – todos os genes de um organismo – em vários estágios de crescimento e desenvolvimento ou em resposta a mudanças ambientais. Estas poderosas novas ferramentas prometem dar inúmeras informações sobre genes e como eles interagem uns com os outros e com o ambiente.

PONTO IMPORTANTE

- Genômica é a subdisciplina da genética devotada a mapeamento, seqüenciamento e análise funcional e comparativa de genomas.

CORRELAÇÃO DOS MAPAS GENÉTICO, CITOLÓGICO E FÍSICO DE CROMOSSOMOS

As localizações cromossômicas de genes e outros marcadores moleculares podem ser mapeadas com base em freqüências de recombinação, posições relativas de características citológicas ou distâncias físicas.

A capacidade dos cientistas de identificar e isolar genes com base em informações sobre sua localização no genoma foi uma das primeiras contribuições importantes da pesquisa genômica. Em princípio, este enfoque, chamado de *clonagem posicional*, pode ser usado para identificar e clonar qualquer gene com um efeito fenotípico conhecido em qualquer espécie. A clonagem posicional tem sido muito usada em várias espécies, incluindo seres humanos. No Cap. 17, consideraremos o uso da clonagem posicional para identificar os genes humanos responsáveis pela doença de Huntington e a fibrose cística.

Como a utilidade da clonagem posicional depende da disponibilidade de mapas detalhados de regiões dos cromossomos onde estão os genes de interesse, importantes esforços têm focado o desenvolvimento de mapas detalhados do genoma humano e dos genomas de importantes organismos modelos, tais como *D. melanogaster*, *C. elegans* e *A. thaliana*. A meta dessa pesquisa é construir mapas genéticos e físicos correlatos com marcadores distribuídos em intervalos relativamente curtos pelo genoma. No caso dos genomas humano e de *Drosophila*, os mapas genéticos e físicos também podem ser correlacionados com mapas citológicos (padrões de bandeamento) dos cromossomos (Fig. 16.2). Discutiremos a construção desses mapas nas seções seguintes deste capítulo.

Lembre que os mapas genéticos (Fig. 16.2, à esquerda) são construídos a partir de freqüências de recombinação, sendo 1 centiMorgan (cM) igual à distância que dá uma freqüência média de recombinação de 1% (Cap. 7). Mapas genéticos com marcadores espaçados em intervalos curtos – mapas genéticos de alta densidade – em geral são construídos usando-se marcadores moleculares tais como fragmentos de restrição de tamanhos diferentes (*polimorfismos de comprimento de fragmentos de restrição*, ou *RFLP*). Mapas citológicos (Fig. 16.2, centro) são baseados em padrões de bandeamento de cromossomos observados ao microscópio após tratamento com vários corantes (Cap. 6). **Mapas físicos** (Fig. 16.2, à direita), tais como os mapas de restrição (Fig. 16.2, em cima, à direita) discutidos no Cap. 15, são baseados nas distâncias moleculares – pares de bases (pb), quilobases (kb, 1.000 pb)

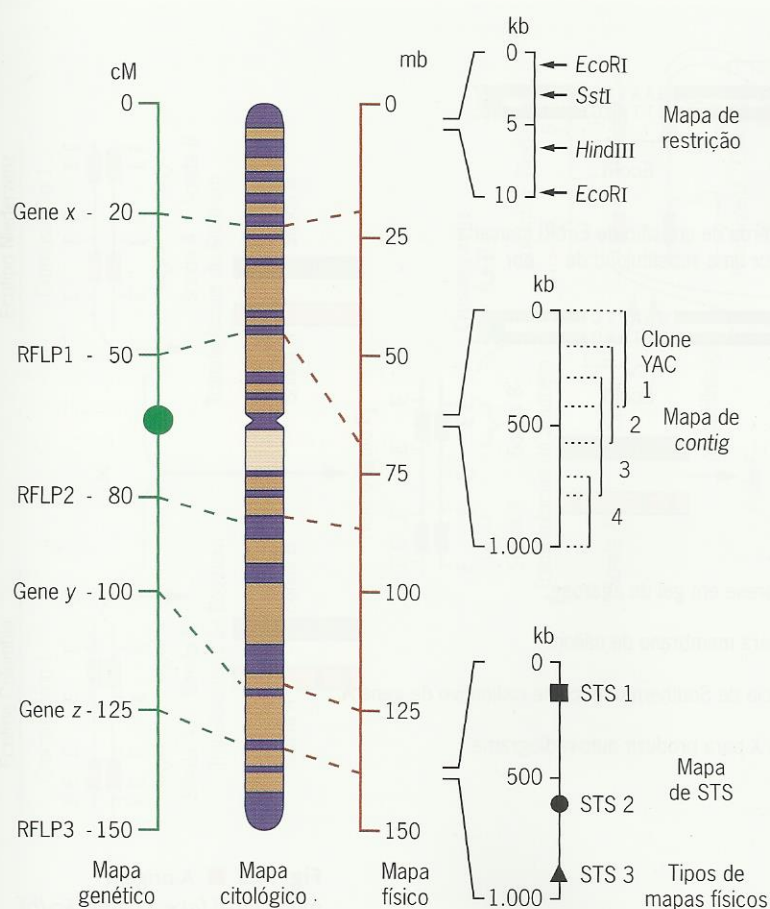


Fig. 16.2 ■ Correlação dos mapas genético, citológico e físico de um cromossomo. As distâncias de mapa genético são baseadas em freqüências de *crossing* e são medidas em porcentagem de recombinação, ou centiMorgans (cM), enquanto as distâncias físicas são medidas em pares de quilobases (kb) ou megabases (mb). Mapas de restrição, mapas de contig e mapas STS (sítios de seqüências marcadas) são descritos no texto.

e megabases (mb, 1 milhão de pb) – que separam sítios em moléculas gigantes de DNA presentes em cromossomos. Os mapas físicos também contêm as localizações de clones genômicos superpostos ou *contigs* (Fig. 16.2, centro, à direita) e seqüências nucleotídicas únicas chamadas *sítios de seqüência marcada*, ou *STS* (Fig. 16.2, embaixo, à direita).

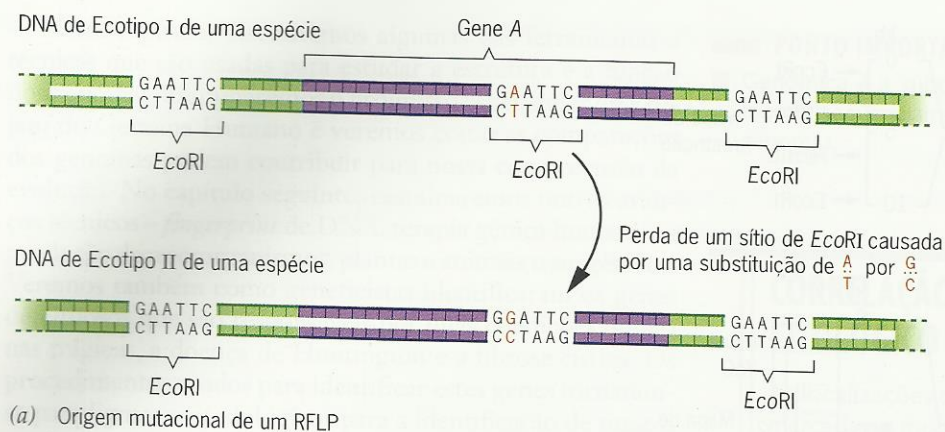
Os mapas físicos de um cromossomo podem estar correlacionados com os mapas genéticos e citológicos de vários modos. Genes que foram clonados podem ser posicionados no mapa citológico por hibridização *in situ* (Cap. 9). Podem ser estabelecidas correlações entre os mapas genético e físico localizando-se clones de genes geneticamente mapeados ou RFLP no mapa físico. Marcadores que são mapeados tanto geneticamente quanto fisicamente são chamados de **marcadores âncoras**; eles ancoram o mapa físico ao mapa genético e vice-versa. Mapas físicos de cromossomos também podem ser correlacionados com mapas genéticos e citológicos usando-se (1) PCR (veja Fig. 15.11) para amplificar seqüências únicas de DNA genômico, geralmente de 200 a 500 pb, (2) transferências de Southern para correlacionar estas seqüências com clones superpostos em mapas físicos e (3) hibridização *in situ* para determinar suas localizações cromossômicas (posições em mapas citológicos). Essas seqüências âncoras únicas curtas são denominadas **sítios de seqüência marcada (STS)**. Outro enfoque usa seqüências

curtas de cDNA (DNA cópia de mRNA), ou **marcações de seqüência expressa (EST)**, como sondas de hibridização para ancorar mapas físicos a mapas de RFLP (mapas genéticos) e mapas citológicos.

Distâncias físicas não se correlacionam diretamente com distâncias de mapa genéticas porque as freqüências de recombinação não são sempre proporcionais às distâncias moleculares. Entretanto, as duas em geral são razoavelmente bem correlacionadas em regiões eucromáticas de cromossomos. Em humanos, 1 cM é equivalente, em média, a cerca de 1 mb de DNA.

Polimorfismos de Comprimento de Fragmentos de Restrição (RFLP) e Mapas de Microssatélite

Quando mutações alteram as seqüências de nucleotídeos em sítios de clivagem de enzimas de restrição, as enzimas não mais os reconhecem (Fig. 16.3a). Outras mutações podem criar novos sítios de restrição. Tais mutações resultam em variações nos tamanhos dos fragmentos de DNA produzidos por digestão com várias enzimas de restrição (Fig. 16.3b). Tais **polimorfismos de comprimento de fragmentos de restrição**, ou **RFLP**, demonstraram-se valiosos na construção de mapas genéticos detalhados para uso na clonagem



- ETAPA 1 Isolar DNA de cada ecotipo.
- ETAPA 2 Digerir DNA com enzima de restrição *EcoRI*.
- ETAPA 3 Separar fragmentos de restrição por eletroforese em gel de agarose.
- ETAPA 4 Transferir fragmentos de DNA de restrição para membrana de náilon.
- ETAPA 5 Hibridizar fragmentos de DNA em transferência de Southern com clone radioativo de gene A.
- ETAPA 6 Lavar transferência e expô-la a filme de raios X para produzir auto-radiograma.

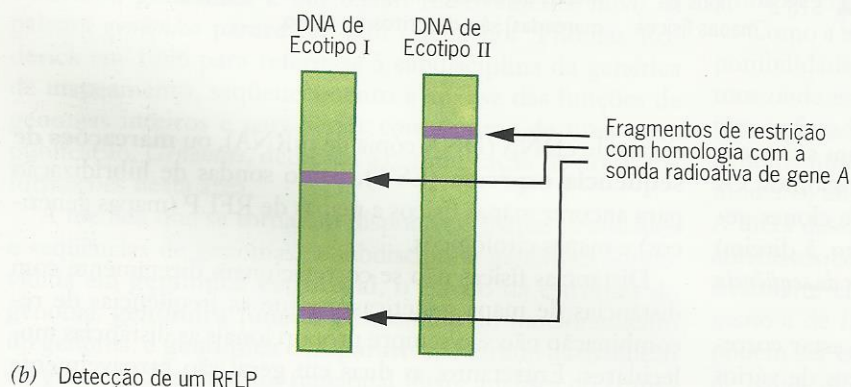


Fig. 16.3 ■ A origem mutacional (a) e a detecção (b) de RFLP em ecotipos diferentes de uma espécie. No exemplo mostrado, uma substituição de pares de bases A:T → G:C resulta na perda da sequência central de reconhecimento de *EcoRI* presente no gene A do DNA do ecotipo I. Esta mutação pode ter ocorrido em ancestral do ecotipo II durante os primeiros estágios de sua divergência do ecotipo I.

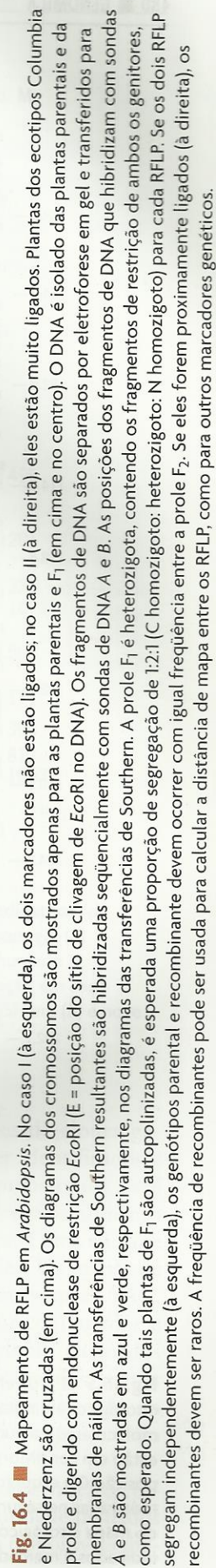
posicional. RFLP são mapeados como outros marcadores genéticos. Eles segregam em cruzamentos como se fossem alelos co-dominantes.

Os DNA de isolados geográficos diferentes, ecotipos diferentes (linhagens adaptadas a condições ambientais diferentes) e linhagens endocruzadas diferentes de uma espécie contêm muitos RFLP que podem ser usados para construir mapas genéticos detalhados. Os DNA de indivíduos diferentes, mesmo parentes, em geral exibem RFLP. Alguns RFLP podem ser visualizados diretamente quando os fragmentos de DNA digeridos são separados por eletroforese em gel de agarose, corados com brometo de etídio e vistos sob luz ultravioleta. Outros RFLP podem ser detectados apenas usando-se cDNA específico ou clones genômicos

como sondas de hibridização radioativas em transferências de Southern genômicas (Fig. 16.3b). Os próprios RFLP são os fenótipos usados para classificar a prole de cruzamentos como parentais ou recombinantes. A Fig. 16.4 ilustra o uso de RFLP como marcadores genéticos em experimentos de mapeamento na planta modelo *Arabidopsis*. Como seu genoma é pequeno (cerca de 10^8 pares de nucleotídeos), *Arabidopsis* é um sistema excelente que permite clonar genes com base em suas localizações.

Os marcadores RFLP são especialmente valiosos no mapeamento de cromossomos de humanos, em que os pesquisadores devem basear-se na segregação de alelos mutantes de ocorrência espontânea em famílias para estimar distâncias de mapa. Mapeamento deste tipo baseado em heredogramas é

Caso II: Marcadores A e B bem ligados



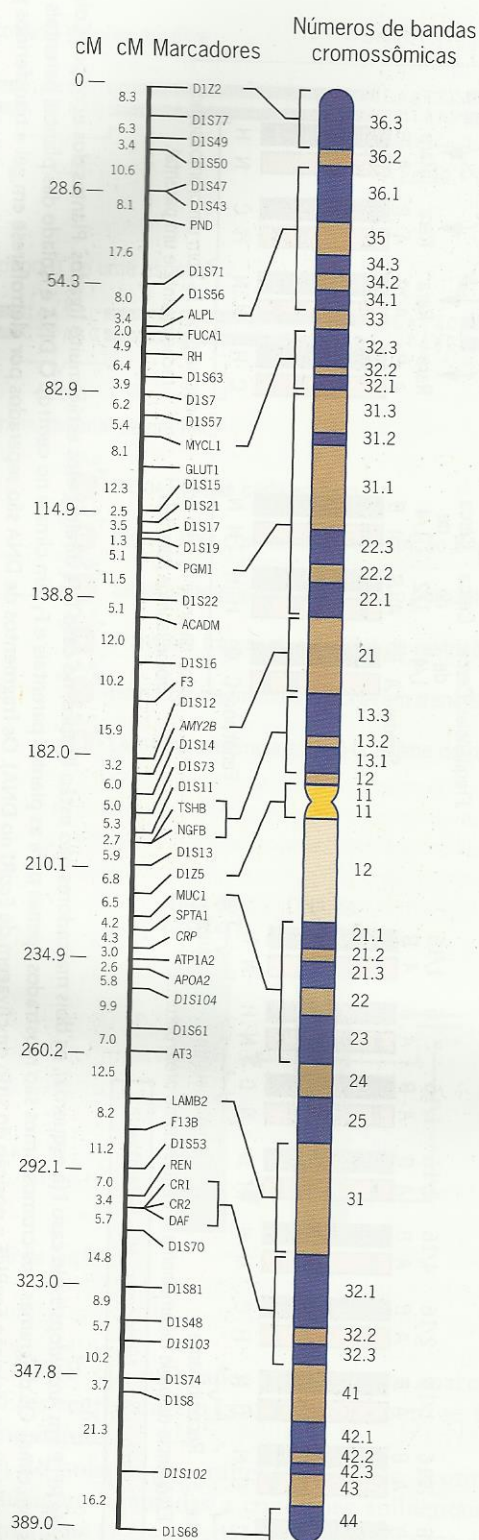


Fig. 16.5 ■ Correlação entre o mapa de RFLP (à esquerda) e o mapa citológico (à direita) do cromossomo humano 1. Os marcadores moleculares e alguns genes são mostrados no centro. As distâncias estão em centiMorgans (cM), com o marcador mais superior na posição 0 à esquerda e as distâncias entre marcadores adjacentes mostradas na segunda coluna da esquerda. Os colchetes à esquerda do mapa citológico mostram os locais cromossômicos dos genes indicados e marcadores moleculares.

feito por comparação das probabilidades de que os marcadores genéticos que segregam no heredograma são não ligados ou são ligados por várias distâncias de mapa. Em 1992, geneticistas usaram este procedimento para construir um mapa inicial com cerca de 2.000 RFLP nos 24 cromossomos humanos. A Fig. 16.5 mostra a correlação entre um mapa de RFLP e o mapa citológico do cromossomo humano 1.

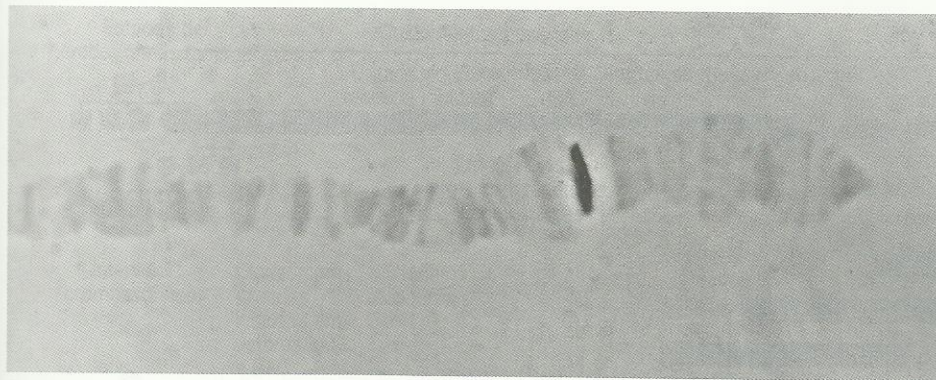
Em humanos, os RFLP mais úteis envolvem seqüências pequenas que estão presentes em repetições em tandem. O número de cópias de cada seqüência presente em determinado sítio em um cromossomo é altamente variável. Tais sítios, chamados de **número variável de repetições em tandem**, ou **VNTR**, são portanto altamente polimórficos. Os VNTR variam de tamanho dos fragmentos não por causa de diferenças nas posições dos sítios de clivagem das enzimas de restrição, mas devido a diferenças no número de cópias da seqüência repetida entre os sítios de restrição.

Microsatélites são uma outra classe de polimorfismos que mostraram ser extremamente valiosos na construção de mapas de alta densidade de cromossomos eucarióticos. Os microsatélites são repetições em tandem polimórficas de seqüências com apenas dois a cinco pares de nucleotídeos de tamanho. Em humanos, as seqüências microsatélites compostas de repetições polimórficas em tandem das seqüências de dinucleotídeos AC/TG (AC em um filamento; TG no filamento complementar) são marcadores especialmente úteis. Em 1996, um grupo de pesquisadores franceses e canadenses publicou um mapa de 5.264 microsatélites AC/TG no genoma humano. Esses microsatélites definiram 2.335 sítios com uma distância média de 1,6 cM ou cerca de 1,6 mb entre marcadores adjacentes.

Em 1997, um grande consórcio internacional usou RFLP para mapear mais de 16.000 genes humanos (EST e genes clonados) e integrou seu mapa ao mapa físico do genoma humano. Neste estudo colaborativo, mais de 20.000 STS foram mapeados em 16.354 *loci* distintos. Esses mapas genéticos compostos primariamente de marcadores RFLP possibilitaram identificar e caracterizar genes mutantes que são responsáveis por muitas doenças humanas (Cap. 17).

Mapas Citogenéticos

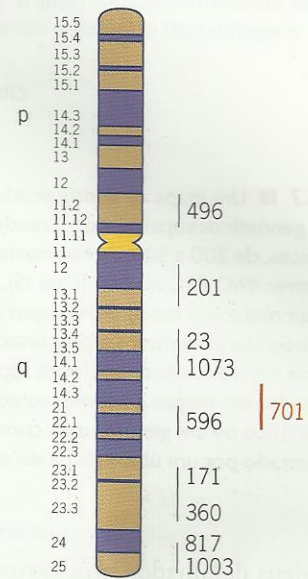
Em algumas espécies, genes e clones podem ser posicionados nos mapas citológicos dos cromossomos por hibridização *in situ* (Cap. 9). Por exemplo, em *Drosophila*, os padrões de bandamento dos cromossomos politênicos gigantes nas glândulas salivares (Cap. 6) fornecem mapas de alta resolução dos cromossomos (Fig. 16.6a). Assim, um clone de conteúdo genético desconhecido pode ser posicionado no mapa citológico com considerável precisão. Em mamíferos, incluindo humanos, hibridização *in situ* com fluorescência (FISH; veja Enfoque Técnico: Hibridização *In Situ*, no Cap. 9) pode ser usada para posicionar clones em cromossomos corados por qualquer dos vários protocolos de bandamento cromossômico. A Fig. 16.6 (b e c) ilustra o uso de FISH para posicionar clones no mapa citológico do cromossomo



(a)



(b)



Cromossomo 11

(c)

Fig. 16.6 ■ Determinação das localizações de genes e clones em cromossomos por hibridização *in situ* usando sondas radioativas (a) ou fluorescentes (b). Mapeamento por hibridização *in situ* (a) do gene *white* de *D. melanogaster* na região 3C2 do cromossomo X. (b) Mapeamento de hibridização *in situ* com fluorescência (FISH) do clone D11RP-701 no intervalo q14.3-22.1 do cromossomo humano 11. Note que estes são cromossomos metafásicos contendo duas cromátides, e, portanto, não existem dois sinais de hibridização em cada cromossomo. (c) Mapa citológico do cromossomo 11 mostrando as localizações do clone (barra vermelha) mapeado por FISH em (b) e nove outros clones (barras pretas) mapeados no cromossomo 11 no mesmo estudo.

humano 11. Se puderem ser identificados RFLP que se superpõem a estes clones, eles podem ser usados como sítios de STS que ancoram o mapa genético do cromossomo 11 ao mapa citológico, produzindo um mapa citogenético. Se os clones podem ser posicionados no mapa físico por experimentos de hibridização de transferência de Southern (Cap. 15), eles também podem ser usados para ligar o mapa físico a mapas genético e citológico do cromossomo.

Mapas Físicos e Bancos de Clones

O procedimento de mapeamento de RFLP tem sido usado para construir mapas genéticos detalhados de cromossomos, os quais, por sua vez, possibilitaram a clonagem posicional. Esses mapas genéticos foram suplementados com mapas físicos de cromossomos. Isolando-se e preparando-

se mapas de restrição de grandes números de clones genômicos, os clones superpostos podem ser identificados e usados para construir mapas físicos de cromossomos e até mesmo genomas inteiros. Em princípio, este procedimento é simples (Fig. 16.7). Entretanto, na prática, é uma enorme tarefa, especialmente para genomas grandes. Os mapas de restrição de grandes clones genômicos em vetores YAC, PAC ou BAC (Cap. 15) são analisados por computador e organizados em grupos superpostos de clones chamados *contigs*. À medida que mais dados são adicionados, *contigs* adjacentes são acrescentados. Quando o mapa físico de um genoma está completo, cada cromossomo é representado por um único mapa de *contig*.

A construção de mapas físicos de genomas inteiros requer que seja pesquisada uma grande quantidade de dados quanto a superposições. Entretanto, estão disponíveis ma-

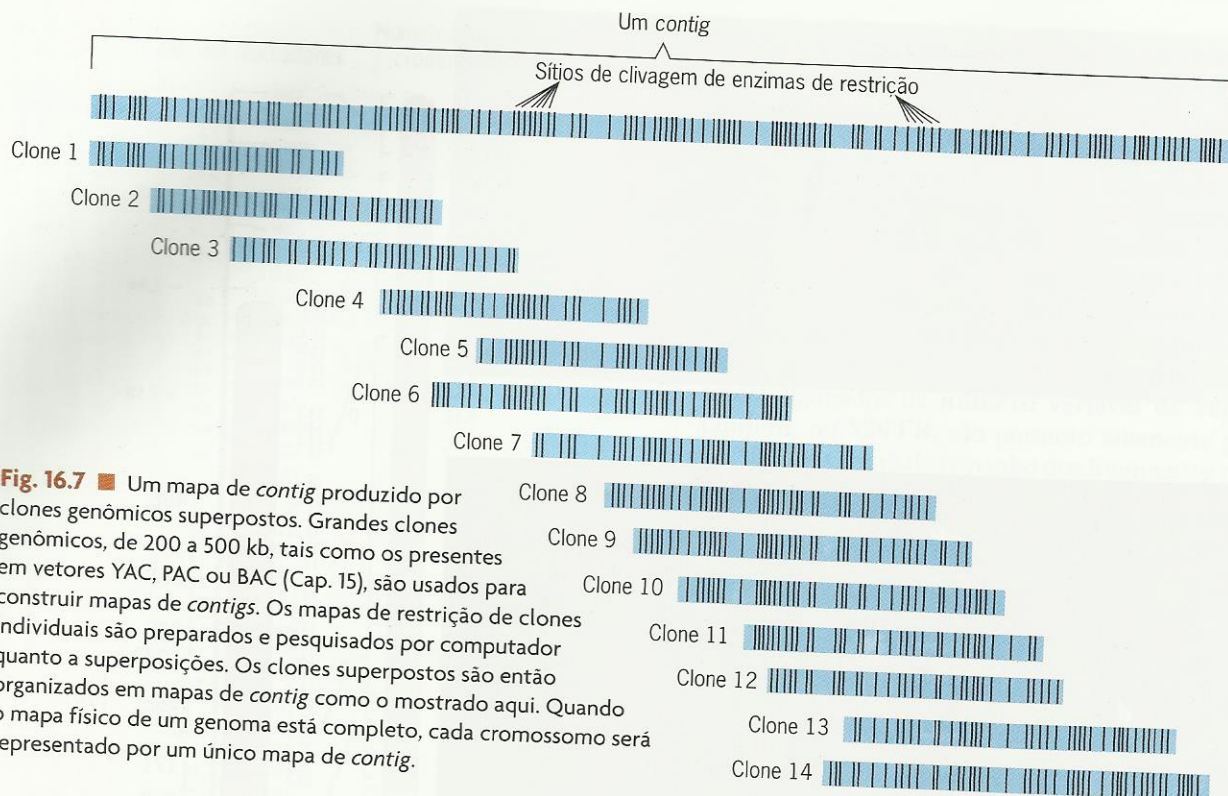


Fig. 16.7 ■ Um mapa de *contig* produzido por clones genômicos superpostos. Grandes clones genômicos, de 200 a 500 kb, tais como os presentes em vetores YAC, PAC ou BAC (Cap. 15), são usados para construir mapas de *contigs*. Os mapas de restrição de clones individuais são preparados e pesquisados por computador quanto a superposições. Os clones superpostos são então organizados em mapas de *contig* como o mostrado aqui. Quando o mapa físico de um genoma está completo, cada cromossomo será representado por um único mapa de *contig*.

pas físicos detalhados para vários genomas, incluindo o de seres humanos, o de *C. elegans*, o de *D. melanogaster* e o de *A. thaliana*. Tais mapas físicos estão sendo usados para preparar bancos de clones que contêm clones catalogados cobrindo coletivamente cromossomos inteiros. Assim, se um pesquisador precisa de um clone de um determinado gene ou segmento de um cromossomo, esse clone já pode ter sido catalogado no banco de clones, ou biblioteca de clones, e estar disponível mediante solicitação. Obviamente, a disponibilidade de tais bancos de clones e mapas físicos correlatos de genomas inteiros estão acelerando muito as pesquisas genéticas. Pesquisar um gene específico com e sem a ajuda de um mapa físico é como procurar um livro em uma enorme biblioteca com e sem um catálogo computadorizado que forneça as localizações dos livros na biblioteca.

PONTOS IMPORTANTES

- Mapas genéticos de cromossomos são baseados em frequências de recombinação entre marcadores.
- Mapas citogenéticos são baseados na localização de marcadores dentro, ou perto, de características citológicas de cromossomos observadas por microscopia.
- Mapas físicos de cromossomos são baseados em distâncias de pares de bases, pares de quilobases ou pares de megabases que separam marcadores.
- Mapas de alta densidade que integram os mapas genético, citológico e físico de cromossomos foram construídos para muitos cromossomos, incluindo todos os cromossomos humanos.

CLONAGEM DE GENES BASEADA EM POSIÇÃO DE MAPA

Mapas genético, citogenético e físico detalhados de cromossomos permitem que cientistas isolem genes por “andar” nos cromossomos ou “saltar” nos cromossomos.

Os primeiros genes eucarióticos a serem clonados foram genes que são expressos em níveis muito altos em tecidos ou células especializadas. Por exemplo, cerca de 90% das proteínas sintetizadas em reticulócitos de mamíferos é hemoglobina. Assim, os mRNA de α - e β -globina podem ser facilmente isolados de reticulócitos e usados na preparação de sondas de cDNA radioativas para triagens de bibliotecas genômicas. Entretanto, a maioria dos genes não são expressos em níveis tão altos em células especializadas. Assim, como são clonados os genes que se expressam em níveis moderados ou baixos? Um enfoque importante é mapear precisamente o gene e pesquisar um clone do gene usando-se procedimentos que dependem de sua localização no genoma. Este enfoque, chamado de **clonagem posicional**, pode ser usado para identificar qualquer gene, fornecendo um mapa adequado da região do cromossomo no qual está localizado.

As etapas na clonagem posicional são ilustradas na Fig. 16.8. O gene é primeiro mapeado em uma região específica de determinado cromossomo por cruzamentos genéticos ou, no caso de humanos, por análise de heredogramas, que geralmente requerem famílias grandes. O gene em seguida

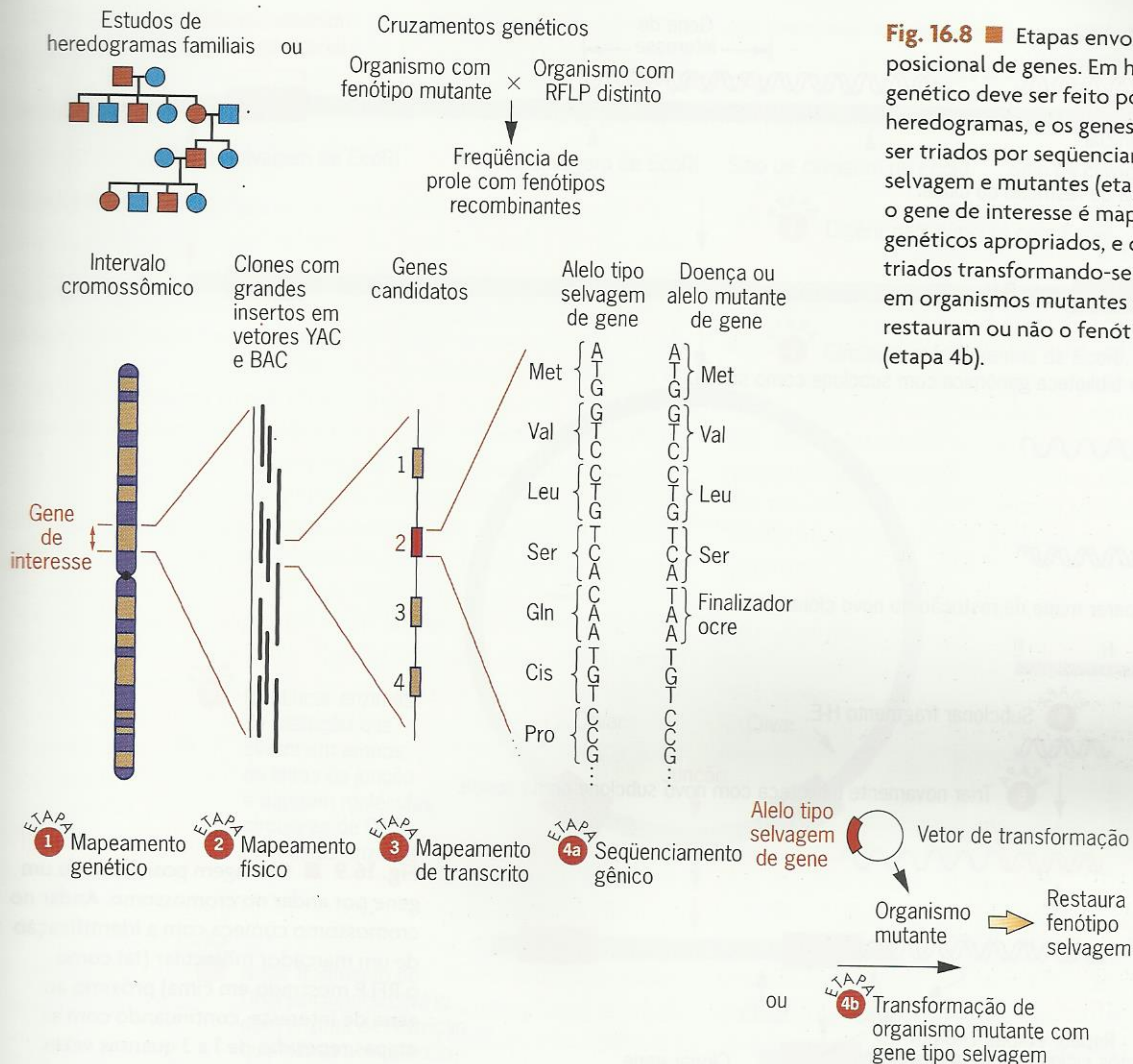


Fig. 16.8 ■ Etapas envolvidas na clonagem posicional de genes. Em humanos, o mapeamento genético deve ser feito por análise de heredogramas, e os genes candidatos devem ser triados por sequenciamento de alelos tipo selvagem e mutantes (etapa 4a). Em outras espécies, o gene de interesse é mapeado por cruzamentos genéticos apropriados, e os genes candidatos são triados transformando-se os alelos tipo selvagem em organismos mutantes e determinando se restauram ou não o fenótipo tipo selvagem (etapa 4b).

é localizado no mapa físico desta região do cromossomo. Genes candidatos no segmento do cromossomo identificados por mapeamento físico são então isolados de indivíduos mutantes e tipo selvagem e sequenciados para identificar mutações que resultariam em perda de função gênica. Como discutiremos no Cap. 17, os genes humanos responsáveis por distúrbios herdados tais como doença de Huntington e fibrose cística foram identificados mediante emprego do enfoque de clonagem posicional. Nas espécies em que é possível a transformação, são introduzidas cópias dos alelos tipo selvagem de genes candidatos em organismos mutantes para determinar se os genes tipo selvagem irão restaurar o fenótipo tipo selvagem. A restauração do fenótipo selvagem em um organismo mutante dá uma forte evidência de que o gene tipo selvagem introduzido é o gene de interesse.

“Andar” no Cromossomo

A clonagem posicional é feita mapeando-se o gene de interesse, identificando um RFLP ou outro marcador molecular próximo ao gene e então “andando” (*walking*) ou “saltando”

(*jumping*) ao longo do cromossomo até encontrar o gene procurado.

Andar no cromossomo é um procedimento que se inicia pela seleção de um marcador molecular (RFLP ou clone gênico conhecido) próximo ao gene de interesse e pelo emprego deste clone como uma sonda de hibridização para triar uma biblioteca de sequências superpostas. Os mapas de restrição são construídos para os clones superpostos identificados na triagem de bibliotecas, e o fragmento de restrição mais distante da sonda original é usado para triar uma segunda biblioteca genômica construída pelo emprego de uma enzima de restrição diferente ou para triar uma biblioteca preparada a partir de uma digestão parcial do DNA genômico. Repetindo este procedimento várias vezes e isolando uma série de clones genômicos superpostos, um pesquisador pode andar ao longo do cromossomo até o gene de interesse (Fig. 16.9). Sem informação sobre a orientação do clone inicial no mapa de ligação, a caminhada inicial terá que continuar em ambas as direções até que seja identificado outro RFLP e seja determinado se o novo RFLP está mais próximo ou mais distante do gene de interesse que o RFLP inicial.

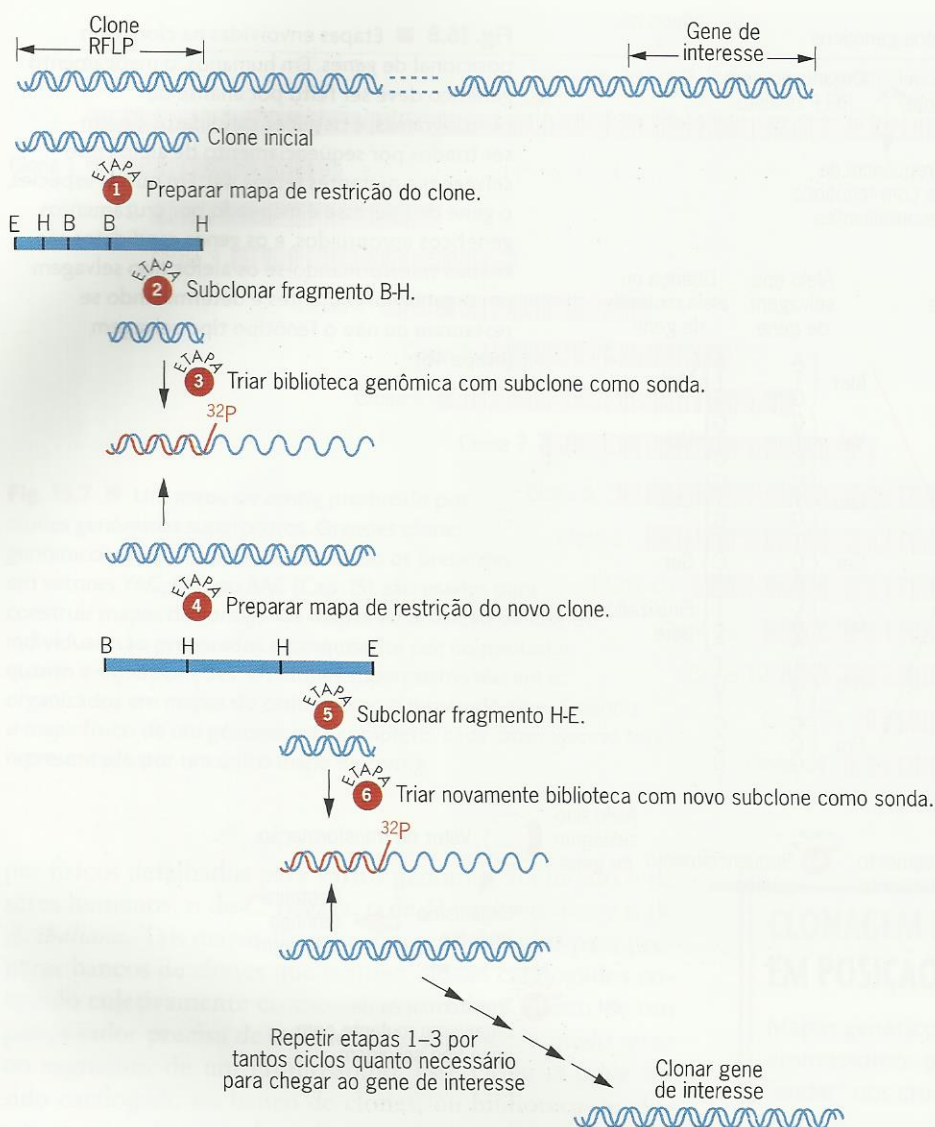


Fig. 16.9 ■ Clonagem posicional de um gene por andar no cromossomo. Andar no cromossomo começa com a identificação de um marcador molecular (tal como o RFLP mostrado em cima) próximo ao gene de interesse, continuando com as etapas repetidas de 1 a 3 quantas vezes sejam necessárias para atingir o gene de interesse (embaixo).

A verificação de que um clone do gene de interesse foi isolado é feita de vários modos. Em organismos experimentais tais como *Drosophila* e *Arabidopsis*, a verificação é obtida introduzindo-se o alelo tipo selvagem do gene em um organismo mutante e mostrando que ele restaura o fenótipo tipo selvagem. Em humanos, a verificação geralmente envolve a determinação das seqüências de nucleotídeos do gene tipo selvagem e vários alelos mutantes e a demonstração de que as seqüências codificantes dos genes mutantes são defeituosas e incapazes de gerar produtos gênicos funcionais.

Andar no cromossomo é tarefa muito difícil em espécies com genomas grandes (a caminhada geralmente é muito longa) e uma abundância de DNA repetitivo disperso (cada seqüência repetida é um potencial bloqueio do caminho). Andar no cromossomo é mais fácil em organismos tais como *A. thaliana* e *C. elegans*, que têm genomas pequenos e pouco DNA repetitivo disperso.

“Saltar” no Cromossomo

Quando a distância do marcador molecular mais próximo do gene de interesse é grande, uma técnica denominada “saltar” no cromossomo pode ser usada para acelerar o que seria uma longa caminhada. Cada salto pode cobrir uma distância de 100 kb ou mais. O procedimento de saltar no cromossomo é ilustrado na Fig. 16.10. Como a caminhada, o salto é iniciado usando-se uma sonda molecular tal como um RFLP como ponto de partida. Entretanto, com os saltos no cromossomo, grandes fragmentos de DNA são preparados por digestão parcial do DNA genômico com uma endonuclease de restrição. Os grandes fragmentos genômicos são então circularizados com DNA-ligase. Uma segunda endonuclease de restrição é usada para remover o fragmento de junção da molécula circular. Esse fragmento de junção conterá ambas as pontas do fragmento longo. Ele pode ser identificado hibridizando-se os fragmentos de DNA em transferências de Southern com a sonda molecular

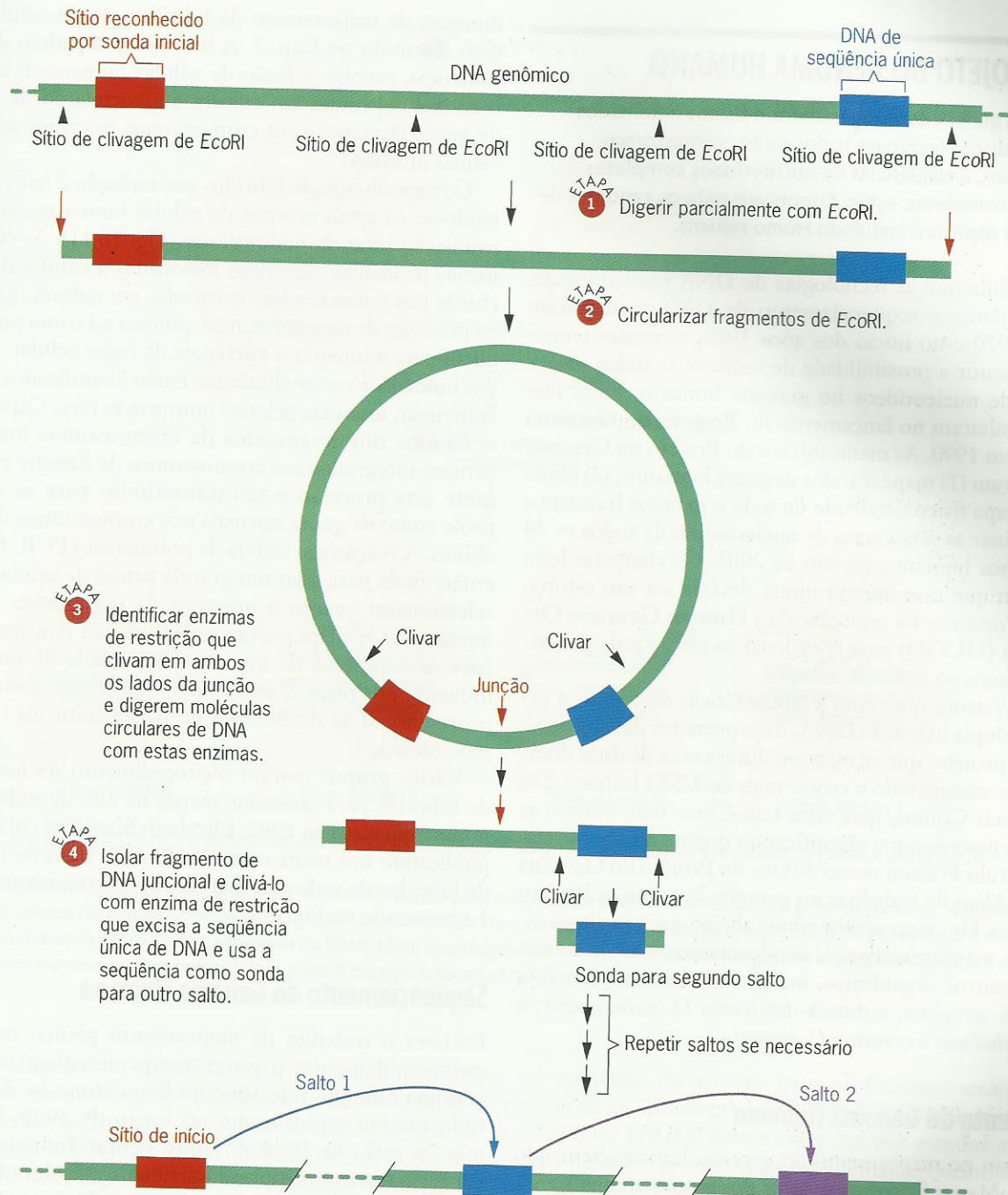


Fig. 16.10 ■ Saltar no cromossomo como um método de atalho para substituir o andar no cromossomo. Este procedimento também pode ser usado para saltar seqüências de DNA repetitivo que bloqueiam o andar no cromossomo.

inicial. É preparado um mapa de restrição do fragmento de junção, e um fragmento de restrição que corresponde à ponta distal do longo fragmento genômico é clonado e usado para iniciar um andar no cromossomo ou um segundo salto no cromossomo. Saltar no cromossomo constitui uma técnica que se mostrou especialmente útil no trabalho com genomas grandes tais como o genoma humano e desempenha um papel importante na identificação do gene humano de fibrose cística (Cap. 17).

■ PONTOS IMPORTANTES

- Mapas genéticos, citogenéticos e físicos detalhados de cromossomos permitem que pesquisadores isolem genes com base em sua localização no genoma.
- Se um marcador molecular tal como um polimorfismo de comprimento de fragmento de restrição (RFLP) é mapeado perto de um gene, o gene pode geralmente ser isolado por procedimentos de caminhar ou saltar no cromossomo.

O PROJETO DO GENOMA HUMANO

Mapas genéticos, citogenéticos e físicos detalhados estão disponíveis para todos os 24 cromossomos humanos, e seqüências de nucleotídeos completas, ou quase completas, estão disponíveis para os genomas de muitas espécies, incluindo *Homo sapiens*.

À medida que as tecnologias de DNA recombinante, clonagem gênica e seqüenciamento de DNA melhoraram nos anos 1970 e no início dos anos 1980, cientistas começaram a discutir a possibilidade de seqüenciar todos os 3×10^9 pares de nucleotídeos no genoma humano. Estas discussões resultaram no lançamento do **Projeto do Genoma Humano** em 1990. As metas iniciais do Projeto do Genoma Humano eram (1) mapear todos os genes humanos, (2) construir um mapa físico detalhado de todo o genoma humano e (3) determinar as seqüências de nucleotídeos de todos os 24 cromossomos humanos no ano de 2005. Os cientistas logo perceberam que essa imensa tarefa deveria ser um esforço mundial. Portanto, foi estabelecida a **Human Genome Organization (HUGO)** para coordenar os esforços de geneticistas humanos ao redor do mundo.

James Watson, que, com Francis Crick, descobriu a estrutura da dupla hélice do DNA, foi o primeiro diretor desse ambicioso projeto, que se estimou durar cerca de duas décadas para ser completado e custar mais de US\$3 bilhões. Em 1993, Francis Collins, que, com Lap-Chee Tsui, liderou as equipes de pesquisa que identificaram o gene da fibrose cística, substituiu Watson como diretor do Projeto do Genoma Humano. Além de trabalhar no genoma humano, o Projeto do Genoma Humano serviu como abrigo para projetos similares de mapeamento e de seqüenciamento de genomas de vários outros organismos, incluindo a bactéria *E. coli*, a levedura *S. cerevisiae*, a mosca-das-frutas *D. melanogaster*, a planta *A. thaliana* e o verme *C. elegans*.

Mapeamento do Genoma Humano

O progresso no mapeamento do genoma humano tem sido excelente. Mapas físicos completos dos cromossomos Y e 21 e mapas detalhados de RFLP do cromossomo X e todos os 22 autossomos foram publicados em 1992. Em 1995, o mapa genético continha marcadores separados por, em média, 200 kb. Um mapa detalhado de microssatélites do genoma humano foi publicado em 1996, e um meticuloso mapa de 16.354 loci distintos foi liberado em 1997. Todos estes mapas se mostraram valiosos para pesquisadores clonarem genes com base em suas localizações no genoma.

Infelizmente, a resolução do mapeamento genético em humanos é muito baixa, na faixa de 1 a 10 mb. A resolução da hibridização *in situ* com fluorescência (FISH) é também de aproximadamente 1 mb. Um mapeamento de melhor resolução (de 50 kb) pode ser obtido por **mapeamento de híbridos por radiação**, que é uma modificação do proce-

dimento de mapeamento de hibridização de células somáticas discutido no Cap. 7. A hibridização padrão de células somáticas envolve a fusão de células humanas e células de roedores cultivadas em cultura e a correlação de produtos de genes humanos com cromossomos humanos retidos nas células híbridas.

O mapeamento de híbridos por radiação é feito fragmentando-se os cromossomos de células humanas com intensa irradiação antes da fusão celular (Fig. 16.11). As células humanas irradiadas são então fusionadas a células de *hamster* chinês (ou outro roedor) cultivadas em cultura, geralmente na presença de uma substância química tal como polietilenglicol para aumentar a eficiência da fusão celular. Os híbridos humano-*hamster* chinês são então identificados mediante cultivo em um meio seletivo apropriado (veja Cap. 7).

Muitos dos fragmentos de cromossomos humanos se tornam integrados aos cromossomos de *hamster* chinês durante este processo e são transmitidos para as células da prole como os genes normais nos cromossomos de *hamster* chinês. A reação em cadeia da polimerase (PCR; Cap. 15) é então usada para triar um grande painel de células híbridas selecionadas quanto à presença de marcadores genéticos humanos. Os mapas cromossômicos são construídos com base na suposição de que a probabilidade de uma quebra induzida por raios X entre dois marcadores é diretamente proporcional às distâncias que os separam no DNA cromossômico.

Vários grupos usaram o procedimento de mapeamento de híbridos para construir mapas de alta densidade do genoma humano. Em 1997, Elizabeth Stewart e colaboradores publicaram um mapa com 10.478 STS baseado em painéis de híbridos de radiação. Seu mapa do cromossomo humano 1 é mostrado na Fig. 16.12.

Seqüenciamento do Genoma Humano

Embora o trabalho de mapeamento gênico tenha avançado rapidamente, o progresso para seqüenciamento do genoma humano inicialmente ficou atrasado. Entretanto, tudo mudou rapidamente no início de 1998. Durante o mês de maio de 1998, J. Craig Venter anunciou que havia formado uma empresa particular, a Celera Genomics, com a meta de seqüenciar o genoma humano em apenas três anos. (Veja detalhes em Um Marco na Genética: Dois Rascunhos da Seqüência do Genoma Humano.) Logo depois, os líderes dos laboratórios públicos de seqüenciamento do Human Genome Project anunciaram que tinham revisto sua programação e planejavam completar a seqüência do genoma humano em 2003, dois anos antes do que havia sido proposto. A partir deste ponto, tudo se acelerou.

A seqüência completa do primeiro cromossomo humano, o pequeno cromossomo 22, foi publicada em dezembro de 1999. A seqüência completa do cromossomo humano 21 surgiu em maio de 2000. Então, por intervenção da White House, Venter, da Celera Genomics, e Fran-

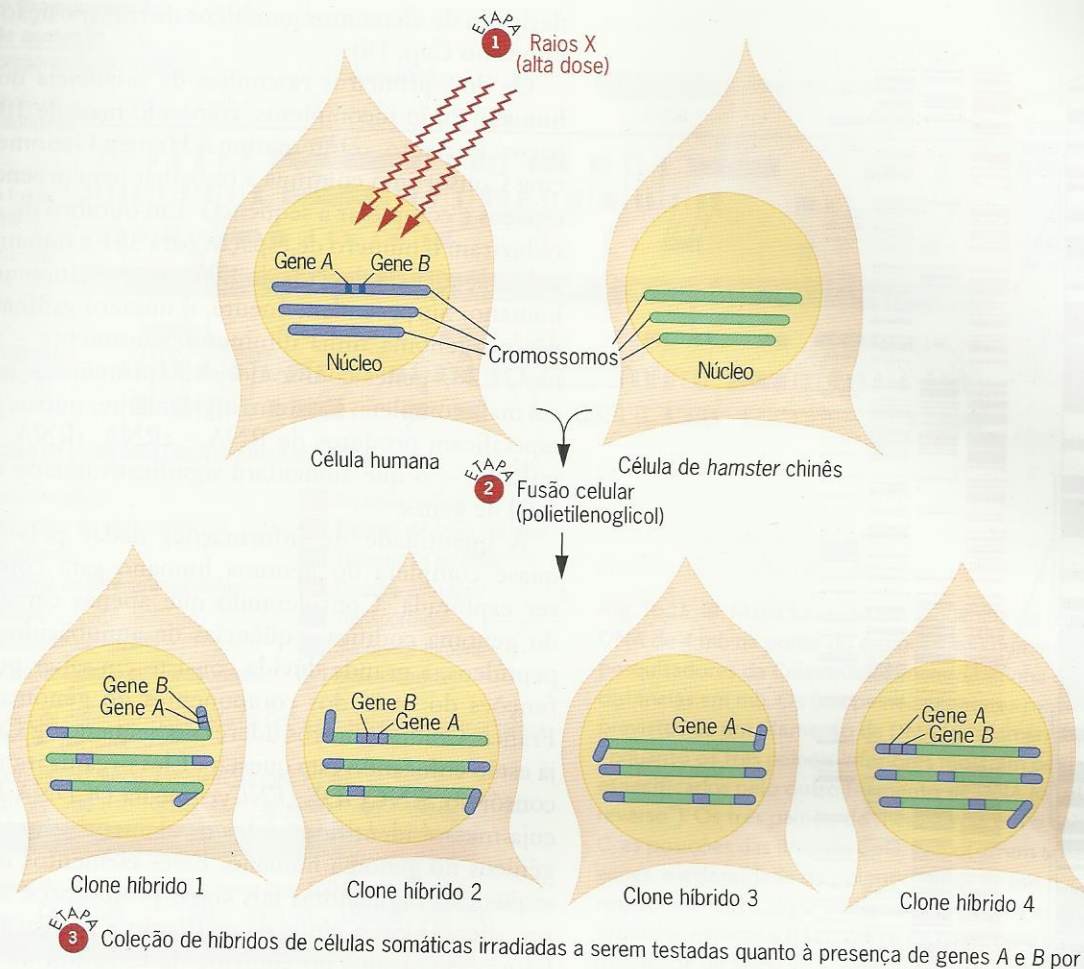


Fig. 16.11 ■ O uso de radiação em híbridos de células somáticas para mapeamento de alta resolução do genoma humano. A lógica subjacente aos mapas de híbridos por radiação é que a probabilidade de uma quebra induzida por raios X entre os genes A e B é diretamente proporcional à distância entre eles na molécula de DNA. Note que os genes A e B permaneceram juntos nos clones híbridos 1, 2 e 4 mas foram separados por uma quebra induzida por raios X durante a formação do clone híbrido 3.

cis Collins, diretor do public Human Genome Project, concordaram em publicar os primeiros rascunhos da sequência do genoma humano ao mesmo tempo. As seqüências da Celera e as públicas foram ambas publicadas em fevereiro de 2001. A Fig. 16.13 mostra um mapa de seqüências de bases de um segmento de 4 mb na ponta do braço curto do cromossomo humano 1. Este mapa ilustra as posições e orientações de genes conhecidos e previstos em uma pequena parte do genoma humano. Mapas similares de todo o genoma humano podem ser vistos no número de 15 de fevereiro de 2001 da *Nature* e de 16 de fevereiro de 2001 da *Science*.

A quantidade de informações nesses primeiros rascunhos do genoma humano é imensa, incluindo a seqüência de mais de 2.650 megabases de pares de DNA (mais de 2.650.000.000 pb). O genoma humano tem mais de 25 vezes o tamanho dos genomas antes seqüenciados de *Drosophila* e *Arabidopsis* e é oito vezes a soma de todos os genomas previamente seqüenciados.

A seqüência do genoma humano forneceu uma surpresa: parece haver apenas cerca de 30.000 a 35.000 genes, e não os 50.000 a 120.000 genes sugeridos por estudos anteriores (veja Enfoque Técnico: Genoma Humano: Quantos Genes?, no Cap. 14). A distribuição de funções para os 26.383 genes previstos pela seqüência da Celera é mostrada na Fig. 16.14. Cerca de 60% das proteínas previstas têm similaridades com proteínas de outras espécies cujos genomas foram seqüenciados (Fig. 16.15). Mais de 40% das proteínas humanas previstas compartilham similaridades com as proteínas de *Drosophila* e *C. elegans*. O quadro é bem diferente de famílias de proteínas bem correlatas, que tendem a desempenhar funções celulares básicas. Apenas 94 das 1.278 famílias de proteínas previstas pela seqüência do genoma humano são específicas de vertebrados. O restante evoluiu de domínios de proteínas em ancestrais distantes, incluindo procariontes e eucariontes unicelulares.

Em média, há um gene a cada 145 kb no genoma humano, embora existam alguns aglomerados de genes al-

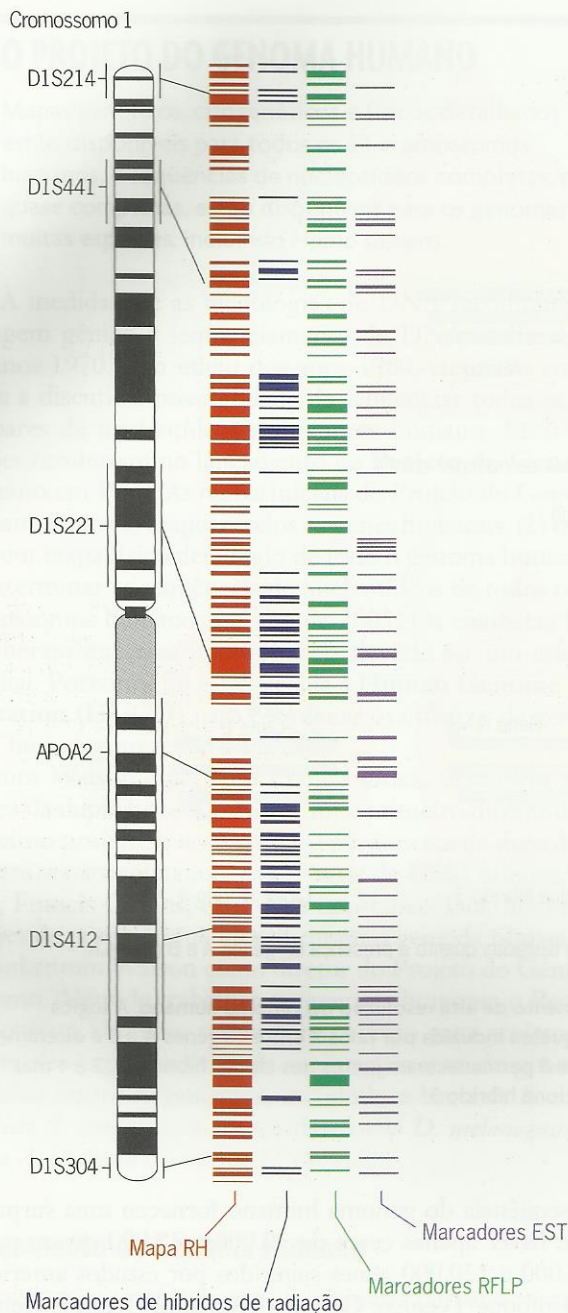


Fig. 16.12 ■ Um mapa híbrido de radiação de alta resolução do cromossomo 1. O mapa citogenético do cromossomo 1 é mostrado à esquerda, juntamente com as localizações de seis marcadores âncoras. À direita do mapa citogenético estão quatro mapas genéticos que mostram as localizações dos marcadores híbridos de radiação (linhas vermelhas), dos marcadores híbridos de radiação de alta confiança (linhas azuis), dos marcadores RFLP (linhas verdes) e das EST (linhas púrpura).

tamente expressos em regiões eucromáticas de cromossomos específicos. O gene humano médio tem cerca de 27.000 pb de tamanho e contém 9 éxons. Éxons constituem apenas 1,1% do genoma, enquanto íntrons constituem 24%, sendo 75% do genoma constituído de DNA intergênico. Do DNA intergênico, pelo menos 44% é

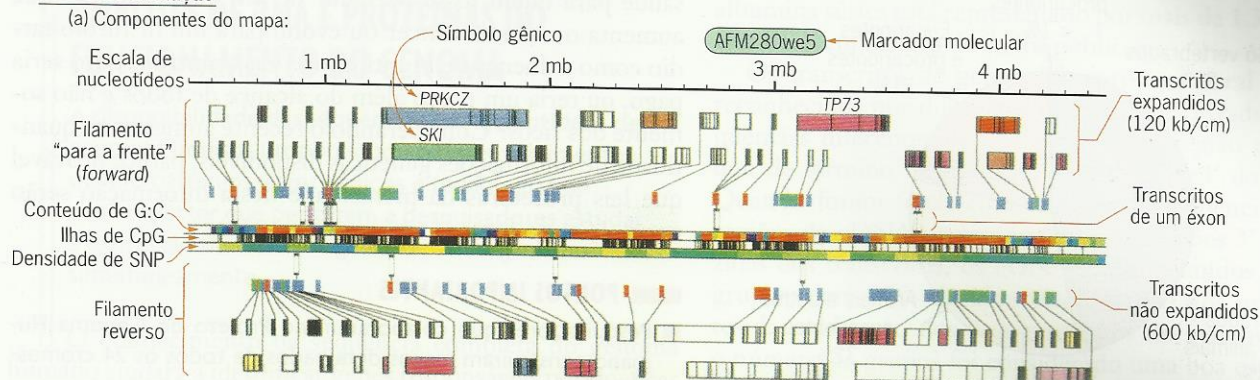
derivado de elementos genéticos de transposição (veja detalhes no Cap. 18).

Os dois primeiros rascunhos da sequência do genoma humano eram incompletos, contendo mais de 100.000 espaços. Portanto, o International Human Genome Sequencing Consortium continua a trabalhar para preencher esses espaços e completar a sequência. Em outubro de 2004, eles reduziram o número de espaços para 341 e tinham completado a sequência de 99% do DNA eucromático no genoma humano. Surpreendentemente, o número estimado de genes no genoma tinha diminuído novamente – para apenas 22.287 genes codificadores de proteínas – na sequência mais completa. Existem, logicamente, outros genes que especificam produtos de RNA – rRNA, tRNA, snRNA e miRNA –, o que aumentará significativamente o número total de genes.

A quantidade de informações dadas pela sequência quase completa do genoma humano está começando a ser explorada. Considerando que apenas cerca de 1,1% do genoma codifica sequências de aminoácidos em polipeptídeos, a grande dúvida consiste em saber quais são as funções do resto dos componentes do genoma humano. Francis Collins e outros líderes do Sequencing Consortium já estão enfocando esta questão. Eles organizaram um novo consórcio, ENCODE (*ENCyclopedia Of DNA Elements*), cuja meta é identificar todos os elementos funcionais não gênicos no genoma humano. Esses elementos irão incluir sequências reguladoras tais como promotores, acentuadores, silenciadores, sítios de metilação e acetilação e outros fatores envolvidos no controle da estrutura da cromatina e da expressão gênica (Fig. 16.16). A identificação e a caracterização funcional desses elementos podem mostrar-se como um desafio maior do que o sequenciamento do genoma humano.

Outro consórcio internacional, o *Human Proteome Organization* (HUPO), foi formado com a meta de determinar as estruturas e funções de todas as proteínas codificadas pelo genoma humano. A despeito do volume de dados fornecidos pela sequência do genoma humano, a dissecação funcional do genoma está apenas começando. Ainda temos um longo caminho a percorrer antes de compreender de fato a estrutura e a função dos 3×10^9 pares de nucleotídeos no genoma humano.

A disponibilidade da sequência do genoma humano levantará inúmeras dúvidas quanto ao uso apropriado deste novo conhecimento. Muitas destas questões enfocam o direito individual a privacidade. Por exemplo, se uma mutação que causa um distúrbio tardio tal como a doença de Huntington (Cap. 17) é descoberta em uma família, quem deve ter acesso a tal informação? Se tais informações estivessem disponíveis para o público, poderia ocorrer uma discriminação generalizada. Empregadores poderiam não contratar membros da família e escolas de medicina poderiam não admitir jovens talentosos em seus cursos de medicina (veja Um Marco na Genética: Mapeamento do Gene para a Doença de

Ponta do Cromossomo 1**Chave de anotação**

(b) Código de cores para função do produto gênico:

Adesão celular	Proteína motora
Regulador de ciclo celular	Ligação de ácido nucleico
Chaperona	Transdução de sinal
Proteína de defesa/imunidade	Proteína estrutural
Enzima	Transportador
Regulador de enzima	Supressor tumoral
União de ligando ou transportador	Desconhecida

(c) Código de cores para conteúdo de G:C e densidade de polimorfismos de um só nucleotídeo (SNP):

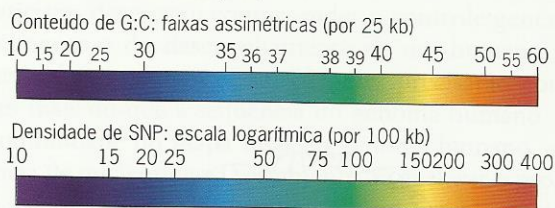


Fig. 16.13 ■ Mapa baseado em sequência de um segmento de DNA de 4 mb na ponta do cromossomo humano 1, montado por pesquisadores da Celera Genomics. (a) A linha de cima fornece as distâncias em mb. Os três painéis seguintes mostram transcritos previstos de um filamento de DNA (o "filamento para a frente"), enquanto os três painéis inferiores mostram os transcritos especificados pelo outro filamento de DNA (o "filamento reverso"). Os três painéis do meio fornecem o conteúdo de G:C, as posições das ilhas de CpG, que ocorrem antecedentes aos genes, e a densidade de polimorfismos de um nucleotídeo (SNP), respectivamente. (b) O código de cores para as funções dos produtos gênicos e (c) o código de cores para o conteúdo de G:C e densidade de SNP.

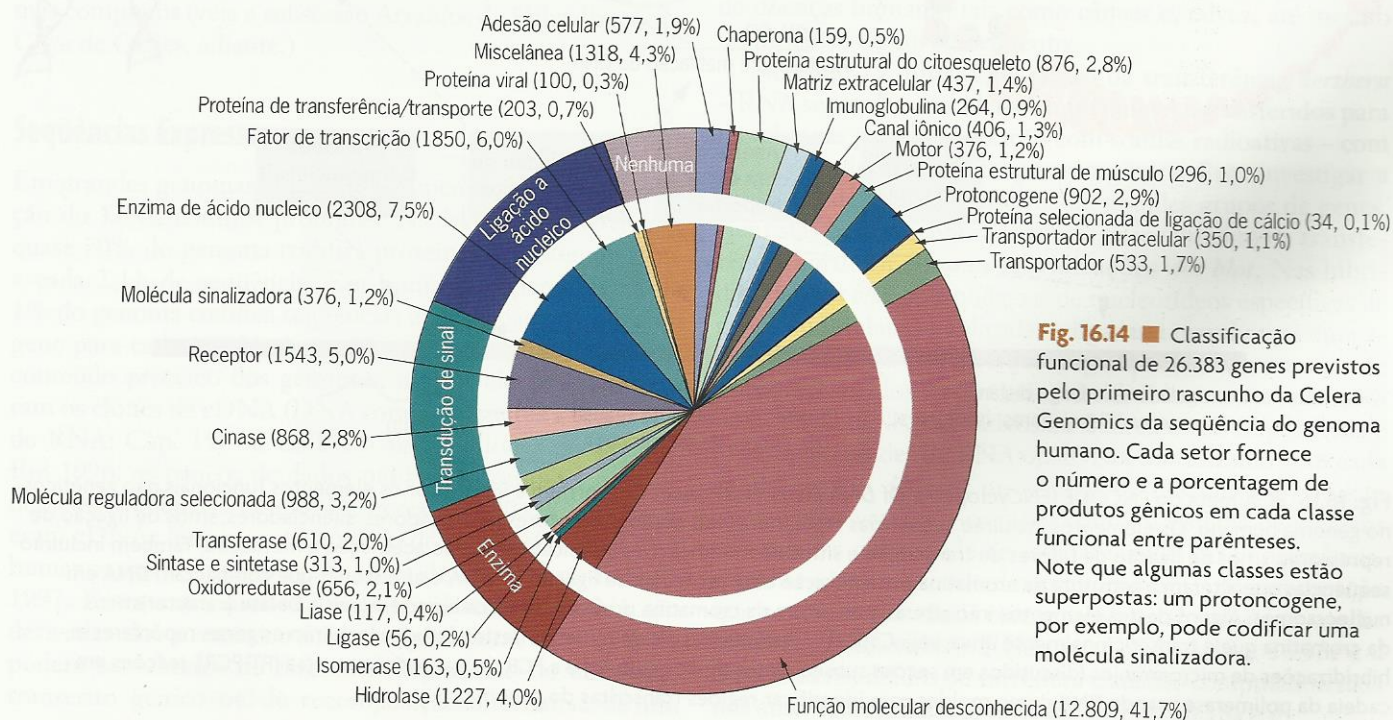


Fig. 16.14 ■ Classificação funcional de 26.383 genes previstos pelo primeiro rascunho da Celera Genomics da sequência do genoma humano. Cada setor fornece o número e a porcentagem de produtos gênicos em cada classe funcional entre parênteses. Note que algumas classes estão superpostas: um protoncogene, por exemplo, pode codificar uma molécula sinalizadora.

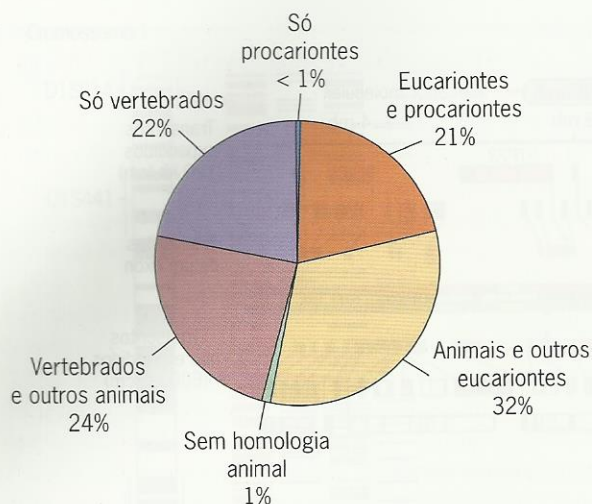


Fig. 16.15 ■ Gráfico mostrando a homologia entre proteínas humanas previstas e proteínas de outras espécies para as quais foram detectados homólogos por pesquisas de computação dos bancos de dados públicos.

Huntington, no Cap. 7). Seguradoras fariam seguros de saúde para quem fosse portador de um gene mutante que aumenta o risco de câncer ou evolui para um distúrbio tardio como a doença de Huntington? Caso sim, o seguro seria pago, ou teria um preço além do alcance de todos e não somente dos ricos? Considerando o recente aumento na quantidade de informações genéticas disponíveis, parece provável que leis protetoras da privacidade desta informação serão necessárias no futuro.

■ PONTOS IMPORTANTES

- Pesquisadores que colaboraram no Projeto do Genoma Humano construíram mapas detalhados de todos os 24 cromossomos humanos.
- Outros participantes do Projeto do Genoma Humano determinaram as seqüências de nucleotídeos completas, ou quase completas, dos genomas de vários organismos modelos importantes.
- Uma seqüência quase completa do DNA eucromático no genoma humano foi publicada em outubro de 2004.

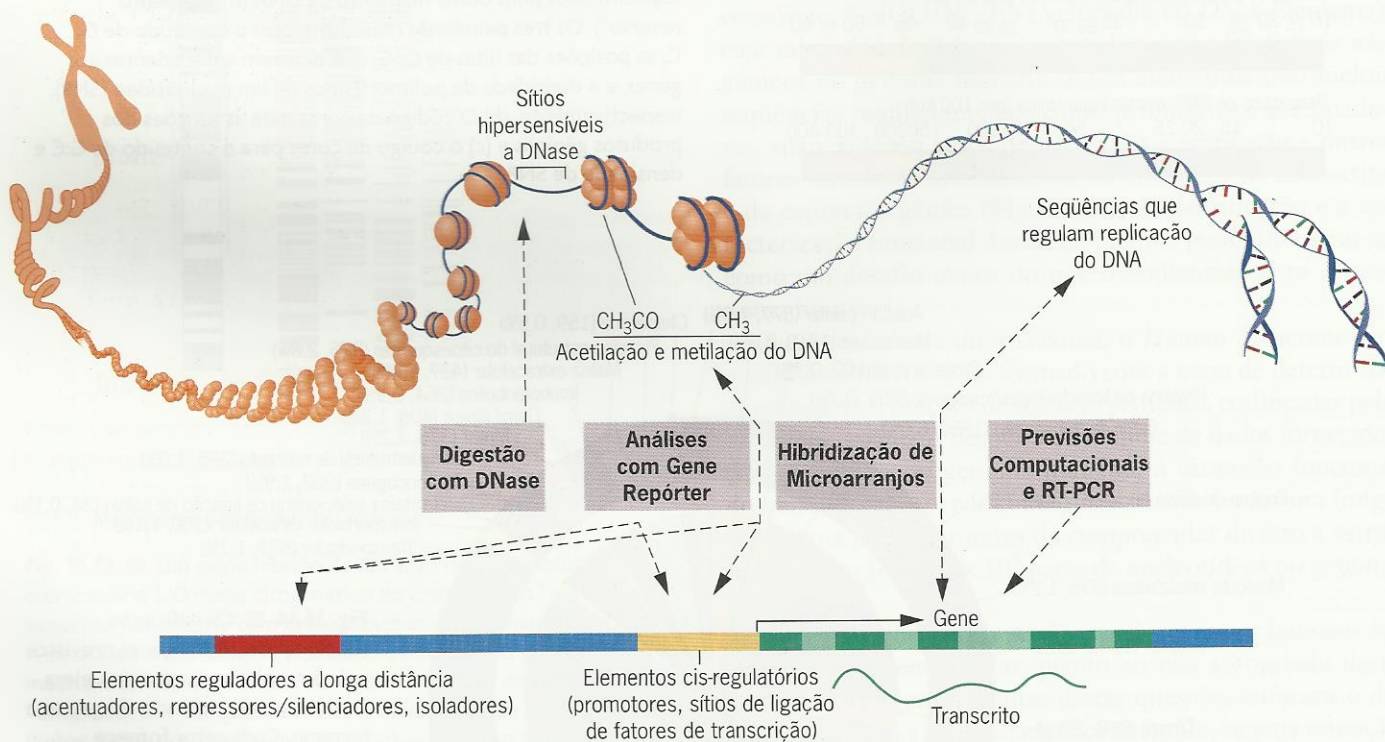


Fig. 16.16 ■ A meta do ENCODE (ENCyclopedia Of DNA Elements) Project Consortium é identificar os elementos funcionais não genéticos no genoma humano. Os elementos incluirão seqüências reguladoras tais como promotores, acentuadores, silenciadores, sítios de ligação de repressores, sítios de ligação de fatores de transcrição e sítios de modificações químicas tais como acetilação e metilação. Também incluirão seqüências que alteram a estrutura da cromatina por interação com proteínas de ligação ao DNA e as histonas que compactam DNA em nucleossomos. Alguns destes elementos irão alterar a estrutura da cromatina produzindo sítios hipersensíveis a DNase (característicos da cromatina que é transcricionalmente ativa; veja Cap. 21). Ferramentas a serem usadas nestes estudos incluirão os genes repórteres e hibridizações de microarranjos (discutidos em seções subseqüentes deste capítulo) e a PCR de transcrição reversa (RT-PCR), reações em cadeia da polimerase usando RNA como moldes para identificar regiões transcritas do genoma.

ANÁLISES DE RNA E PROTEÍNAS DO FUNCIONAMENTO DO GENOMA

A disponibilidade das seqüências de nucleotídeos de genomas inteiros proporcionou o desenvolvimento de microarranjos, *chips* gênicos e tecnologias de gene repórter que permitem a pesquisadores estudar a expressão de todos os genes de um organismo simultaneamente.

O conhecimento da seqüência completa do genoma humano ajudará a identificar genes responsáveis por doenças humanas e deverá resultar em terapias gênicas bem-sucedidas para algumas destas doenças. Entretanto, isto não nos diz o que tais genes fazem ou como eles controlam processos biológicos. De fato, por si só a seqüência de nucleotídeos de um gene, um cromossomo, ou todo um genoma não é informativa. Apenas quando suplementada com informações sobre suas funções é que as seqüências se tornam verdadeiramente significativas. Assim, a informação sobre as funções de seqüências de nucleotídeos ainda deve ser obtida por estudos genéticos tradicionais e por análises moleculares. Se geneticistas desejarem compreender o controle genético do crescimento e do desenvolvimento de um humano adulto a partir de um zigoto (Cap. 22), eles precisarão conhecer muito mais do que a seqüência do genoma humano. Mas a disponibilidade do mapa final do genoma humano, sua seqüência de nucleotídeos, certamente irá acelerar o progresso para a compreensão dos programas de expressão gênica que controlam a morfogênese. O desenvolvimento de novas tecnologias tais como hibridizações de "gene chip" é destinado a tirar proveito da disponibilidade de seqüências de genomas completos (veja a subseção Arranjos de Hibridizações e *Chips* de Genes, adiante.)

Seqüências Expressas

Em grandes genomas eucarióticos, apenas pequena proporção do DNA codifica proteínas. Na levedura *S. cerevisiae*, quase 70% do genoma codifica proteínas, e existe um gene a cada 2 kb de seqüência. Em humanos, apenas cerca de 1% do genoma codifica seqüências de aminoácidos, e há um gene para cada 130 kb de seqüência. Assim, para enfocar o conteúdo protéico dos genomas, muitos cientistas analisaram os clones de cDNA (DNA complementares a moléculas de RNA; Cap. 15) ou EST em vez de clones genômicos. Em 1996, os bancos de dados públicos continham mais de 600.000 seqüências de cDNA, cerca de 450.000 das quais eram cDNAs humanos. O número de seqüências de cDNA humanos quase duplicou para cerca de 800.000 no final de 1997. Entretanto, muitas dessas seqüências de cDNA são derivadas dos mesmos transcritos gênicos. Vários cDNAs podem ser obtidos de diferentes segmentos de um único transcrito gênico ou de recomposição alternativa de um

transcrito gênico. Por exemplo, o gene humano que codifica albumina sérica está representado por mais de 1.300 seqüências EST em bancos de dados públicos.

Os transcritos de genes diferentes em geral podem ser reconhecidos por diferentes regiões 3' não traduzidas – seqüências diferentes de nucleotídeos na região entre o códon de término de tradução 3' e a ponta 3' do transcrito. Quando foram feitas comparações de seqüências entre as regiões de cDNA correspondentes às regiões 3' não traduzidas dos transcritos, os cDNA foram reunidos em 49.625 grupos (com 97% de identidade dentro dos grupos). Antes da publicação dos dois rascunhos do genoma humano, o número de grupos foi considerado uma boa estimativa do número de genes distintos. Dos grupos de seqüências, 4.563 correspondem a genes humanos conhecidos. Um problema com a estimativa do número de genes de seqüências EST é que as EST podem ser derivadas de regiões não superpostas de um transcrito gênico. Em qualquer caso, agora parece claro que o número de genes humanos estimado com base em EST era alto demais.

Arranjos de Hibridizações e *Chips* de Genes

Considerando a seqüência de um genoma inteiro, geneticistas podem começar imediatamente a estudar a expressão de cada gene no organismo. Podem ser sintetizadas sondas oligonucleotídicas de hibridização que são complementares a segmentos dos transcritos de cada ORF ou PCR pode ser empregada para fazer milhões de cópias de cada gene em um genoma. Assim, cientistas podem monitorar mudanças na expressão total do genoma ao longo do tempo por meio de desenvolvimento ou em resposta a mudanças no ambiente. Tal conhecimento deve mostrar-se valioso na compreensão de doenças humanas tais como câncer e, talvez, até mesmo do processo de envelhecimento.

No Cap. 15, discutimos o uso da transferência *northern* – RNA separados por eletroforese em gel e transferidos para membranas para hibridização com sondas radioativas – com o objetivo de estudar a expressão gênica. Para investigar a expressão de genomas inteiros ou grandes grupos de genes, cientistas basicamente reverterem o procedimento da transferência *northern* e fazem **hibridizações dot blot**. Nas hibridizações *dot blot*, as seqüências de nucleotídeos específicas de genes ou sondas são aplicadas e ligadas a membranas em padrões específicos, ou arranjos (Fig. 16.17), e tais arranjos de sondas (chamados *dot blots*) são hibridizados com preparações de RNA ou DNA radioativos ou fluorescentes (Fig. 16.18a e b). As quantidades de RNA ou cDNA hibridizadas com cada sonda em um *dot blot*, ou arranjo de hibridização, podem ser medidas varrendo-se o *blot* com um sistema de imagens que mede a quantidade de radioatividade, ou fluorescência, e analisando-se os resultados com programas de computação que comparam os sinais com aqueles produzidos por sondas controles conhecidas e RNA ou cDNA. A Fig. 16.18c e d mostra os resultados de varredura e análise computadorizada dos auto-radiogramas mostrados na Fig. 16.18a e b.

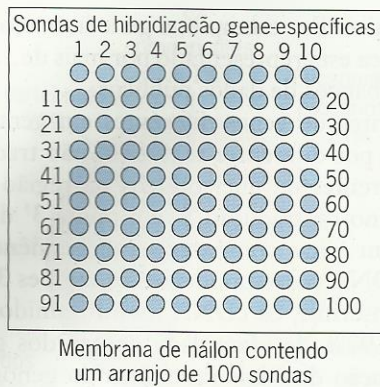
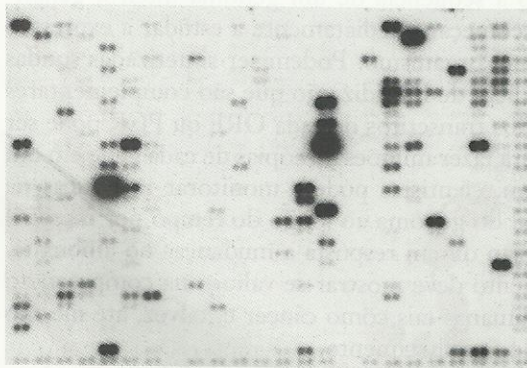


Fig. 16.17 ■ *Dot blot* ou análise de hibridização de arranjos da expressão gênica. Sondas de hibridização específicas de genes são aplicadas a uma membrana com a ajuda de uma filtração a vácuo. A membrana é então colocada em uma solução de hibridização contendo RNA e cDNA marcados com um isótopo radioativo ou um corante fluorescente. Após hibridização, o RNA marcado ligado a cada sonda é ou visualizado por auto-radiografia (veja Fig. 16.18) ou quantificado por densitometria ou escaneadores que medem a quantidade de radioatividade ou fluorescência.

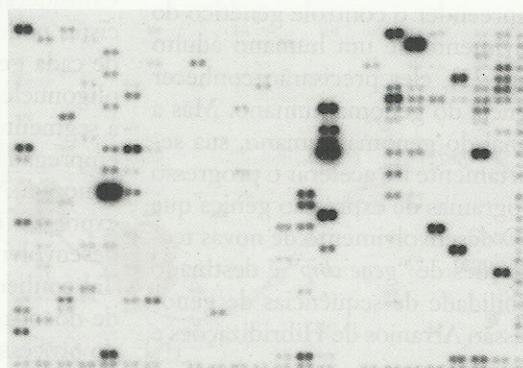
Os arranjos de hibridização mostrados na Fig. 16.18 contêm 588 sondas EST humanas. Entretanto, novas tecnologias agora permitem que cientistas produzam microarranjos que contêm milhares de sondas de hibridização em uma única membrana ou outro suporte sólido. Os *chips* gênicos mostrados na Fig. 16.19 podem conter mais de 10.000 sondas de oligonucleotídeos em uma placa de silicone de apenas alguns centímetros quadrados. Microarranjos de oligonucleotídeos ou sondas de cDNA são produzidos de vários modos: (1) micro-síntese de oligonucleotídeos *in situ* (no *chip*), (2) sondas de oligonucleotídeos pré-fabricadas em suportes sólidos e (3) pontos de fragmentos de DNA ou cDNA nos suportes. As sondas nos microarranjos são então hibridizadas com amostras de cDNA ou RNA marcadas (geralmente com uma marcação fluorescente). A quantidade de RNA ou cDNA hibridizado a cada sonda em um microarranjo é quantificada mediante o emprego de *scanners* sofisticados com resolução micrométrica e um programa de computação apropriado.

Os projetos de seqüenciamento de genoma e as tecnologias de hibridização de microarranjos envolvem a nova disciplina da genômica funcional, que enfoca a expressão de

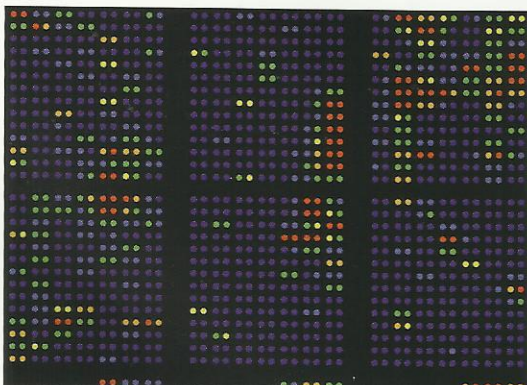
Auto-radiogramas



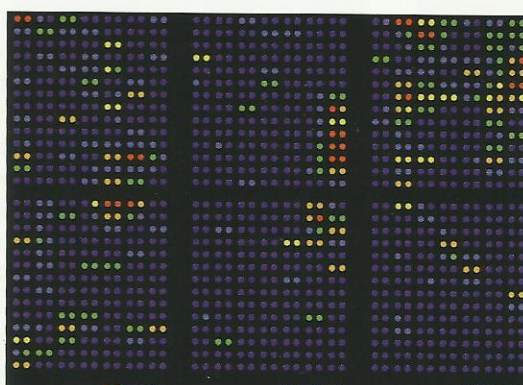
(a) Células cancerosas não tratadas



(b) Células cancerosas tratadas com agente quimioterápico



(c) Análise computadorizada de auto-radiograma em (a)



(d) Análise computadorizada de auto-radiograma em (b)

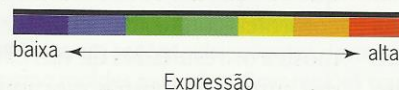


Fig. 16.18 ■ O uso de hibridizações *dot blot*, ou arranjo, para investigar mudanças nos níveis de expressão gênica. Os auto-radiogramas (em cima) mostram os resultados de hibridizações de arranjos de 588 EST humanas com cDNA marcados com ^{32}P preparados de RNA poli(A) isolados de células de câncer humano em cultura, seja (a) não tratadas ou (b) tratadas com um agente quimioterápico. As fotos mostradas em (c) e (d) foram produzidas usando-se um *scanner* para medir as intensidades dos sinais de hibridização nos auto-radiogramas em (a) e (b), respectivamente, e convertendo-os em imagens visuais com programas apropriados de computação. Mudanças nos níveis de expressão gênica induzidas pelo agente quimioterápico podem ser detectadas comparando-se os dois arranjos.

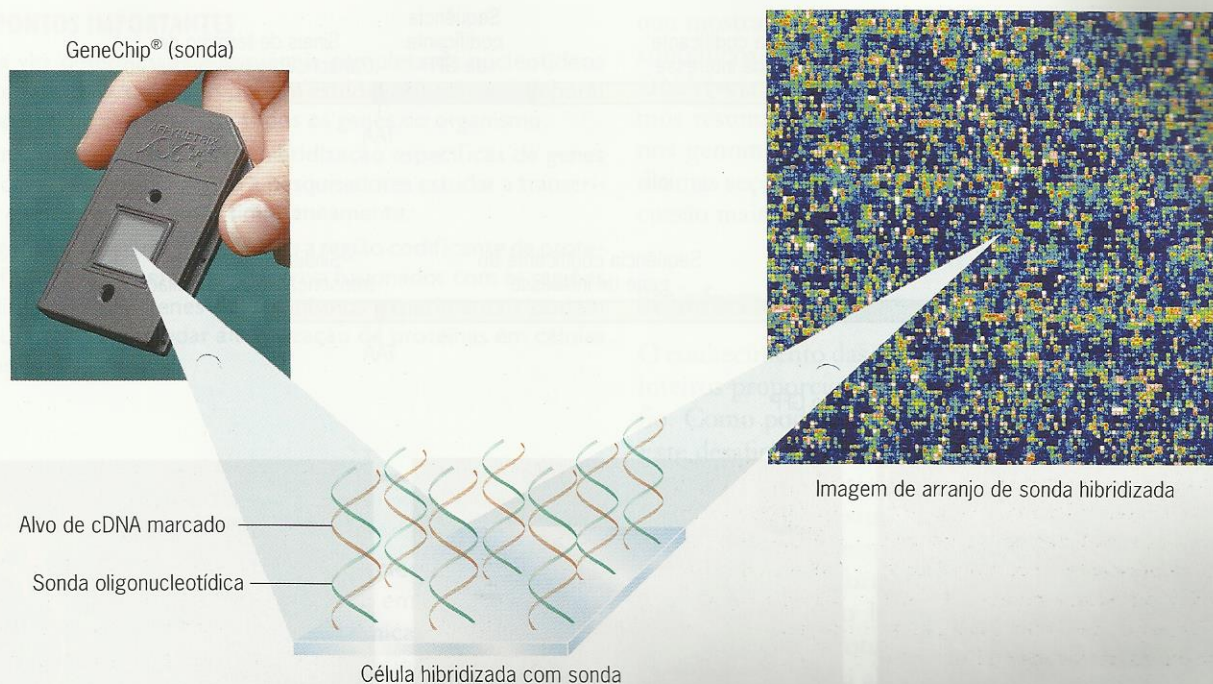


Fig. 16.19 ■ Foto de um *chip* gênico (em cima, à esquerda) e uma ilustração do uso do *chip* gênico para analisar a expressão de todos os genes de um organismo simultaneamente. Estes *chips* gênicos contêm milhares de sondas de hibridização de oligonucleotídeos. *Chips* gênicos permitem que pesquisadores detectem os transcritos de milhares de genes em um experimento.

genomas inteiros. Entretanto, alguns geneticistas dizem que isto tem sido a meta da ciência da genética desde sua concepção. À medida que o conhecimento no campo avançou, geneticistas têm sido capazes de estudar a expressão de cada vez mais genes. Agora, pela primeira vez, eles podem estudar a expressão de todos os genes de um organismo simultaneamente. Microarranjos de sondas que permitem a geneticistas analisar a expressão de quase 6.000 genes de leveduras estão disponíveis, e *chips* de DNA que permitem a cientistas estudar a expressão de aproximadamente 14.000 genes de *Drosophila melanogaster*, os cerca de 26.000 genes de *Arabidopsis thaliana* e os aproximadamente 23.000 humanos estão agora à disposição da comunidade de pesquisadores. A capacidade de analisar a expressão de genomas inteiros acentuará a atual explosão de novas informações na biologia e acabará por proporcionar compreensão do processo normal de desenvolvimento humano e das causas de pelo menos algumas doenças humanas.

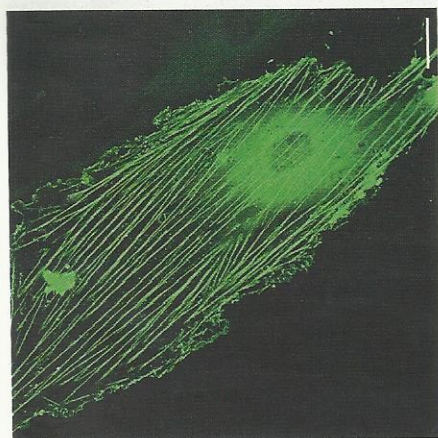
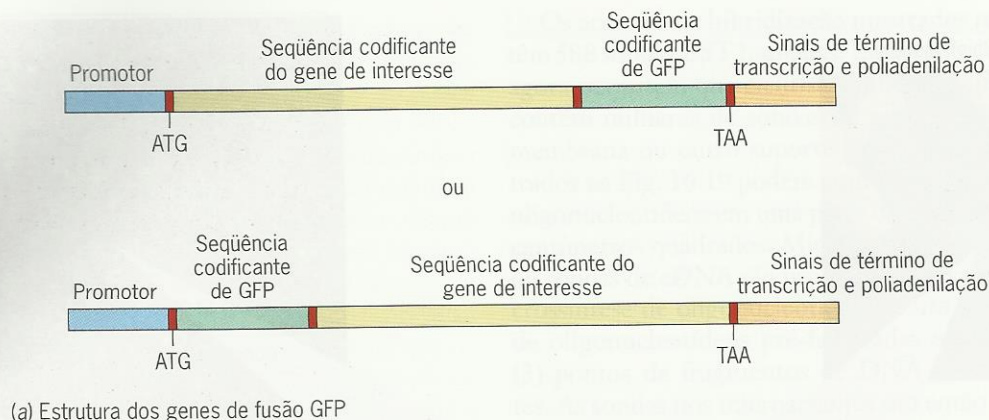
Uso de Proteína Fluorescente Verde como Repórter de Síntese de Proteínas

Arranjos de hibridizações e *chips* gênicos podem ser usados para determinar se genes são transcritos, mas não dão informações sobre a tradução dos transcritos gênicos. Assim, biólogos em geral usam anticorpos para detectar os produtos protéicos de genes de interesse. Transferências

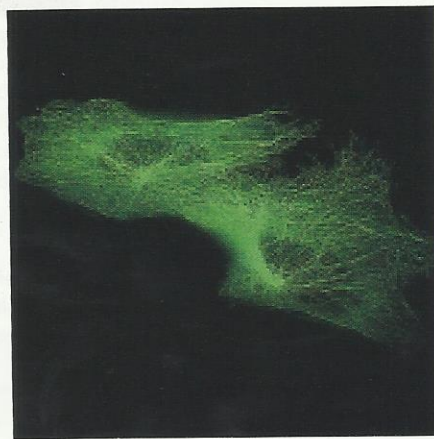
western são usadas para detectar proteínas separadas por eletroforese (Cap. 15), e anticorpos acoplados a compostos fluorescentes são usados para detectar a localização de proteínas *in vivo*. Entretanto, ambos os enfoques só fornecem análise de uma proteína em uma célula, tecido ou organismo.

A descoberta de uma proteína fluorescente de ocorrência natural, a **proteína fluorescente verde (GFP)** da água-viva *Aequorea victoria*, forneceu uma ferramenta importante que pode ser usada para estudar a expressão gênica do nível de proteína. A GFP agora está sendo usada para monitorar a síntese e a localização de proteínas específicas em uma grande variedade de células vivas. Estes estudos envolvem a construção de genes de fusão que contêm as seqüências de nucleotídeos codificantes de GFP, acoplada à seqüência de nucleotídeos codificante da proteína de interesse; a introdução do gene quimérico em células por transformação; e o estudo da fluorescência da proteína de fusão em células transgênicas expostas a luz azul ou UV (Fig. 16.20). Como GFP é uma proteína pequena, ela pode ser acoplada a proteínas sem interferir em sua atividade ou interação com outros componentes celulares.

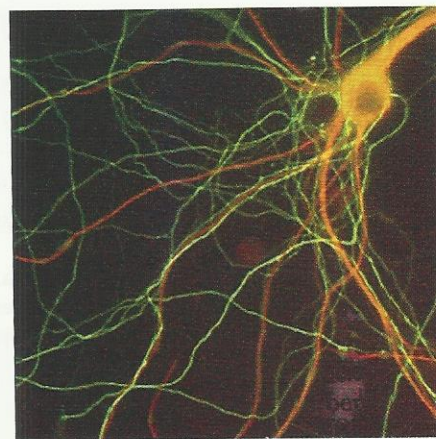
Como o nome indica, GFP fluoresce em verde-claro quando exposta à luz azul ou ultravioleta. O cromóforo da GFP é produzido por ciclização pós-traducional e oxidação de um tripeptídeo codificado serina/tirosina/glicina. Este cromóforo é amplamente protegido de efeitos de íons e solventes por enclausuramento tipo barril da proteína final.



(b) Actina marcada com GFP



(c) Tubulina marcada com GFP



(d) MAP2 marcado com GFP (vermelho) mais tau marcado com GFP (verde)

Fig. 16.20 ■ Uso da proteína fluorescente verde (GFP) da água-viva para estudar a localização de proteínas em células vivas. (a) Estrutura de genes de fusão GFP. A sequência codificante de GFP pode ser colocada em uma das pontas de um gene de interesse ou em posições internas. (b–d) Localização por imunofluorescência de proteínas marcadas com GFP: (b) actina de músculo liso em uma célula fibroblasto; (c) a proteína tubulina da estrutura de microtúbulos em células de ovário de *hamster* chinês; e (d) marcação dupla de duas proteínas de ligação de microtúbulos, MAP2 marcada com GFP emissora de luz azul e tau marcada com GFP emissora de luz verde, em um neurônio de rato. Com os filtros de luz usados para microscopia, MAP2 e tau apresentam-se em vermelho e verde, respectivamente.

Ao contrário de outras proteínas bioluminescentes, a GFP não requer a adição de substratos, cofatores ou nenhuma outra substância para fluorescer quando exposta a luz azul ou UV. Assim, a GFP pode ser usada para estudar a expressão gênica em células vivas e para estudar a localização de proteínas e o movimento em células com o tempo. Mutando o gene da GFP, biólogos moleculares produziram formas variantes de GFP que emitem luz azul ou amarela, variantes que fluorescem até 35 vezes mais intensamente que a GFP tipo selvagem, e variantes cuja fluorescência depende do pH do microambiente. Essas variantes da GFP podem ser usadas para estudo da síntese e da localização intracelular de duas ou mais proteínas simultaneamente (Fig. 16.20d).

Alguns geneticistas estão usando fusões GFP para estudar mudanças na expressão de todos os genes que codificam proteínas envolvidas em determinada via metabólica em res-

posta a tratamento de células ou tecidos com uma droga específica ou agente potencialmente terapêutico. Eles construíram um conjunto inteiro de genes quiméricos contendo a região codificante GFP fusionada a regiões codificantes de outros genes, introduziram tais genes em células hospedeiras e monitoraram sua expressão quantificando a fluorescência de proteínas de fusão separadas por eletroforese ou outras técnicas. Estão sendo desenvolvidas tecnologias que permitirão a cientistas observar mudanças nos níveis de grandes arranjos de proteínas de fusão GFP por eletroforese de capilaridade (eletroforese feita em pequenos tubos capilares), monitorados por microfotodetectores sensíveis e programas de computação sofisticados. Em um futuro não muito distante, a genômica funcional pode envolver o uso de *chips* de DNA para detectar transcritos gênicos e “*chips* de proteína” para detectar os polipeptídeos codificados por estes transcritos.

PONTOS IMPORTANTES

- Uma vez determinada a sequência completa de nucleotídeos de um genoma, cientistas podem estudar os padrões temporal e espacial de expressão de todos os genes do organismo.
- Microarranjos de sondas de hibridização específicas de genes em *chips* gênicos permitem a pesquisadores estudar a transcrição de milhares de genes simultaneamente.
- Os genes quiméricos que contêm a região codificante da proteína fluorescente verde da água-viva fusionados com as regiões codificantes dos genes de organismos experimentais podem ser usados para estudar a localização de proteínas em células vivas.

GENÔMICA COMPARATIVA

Comparações das sequências de nucleotídeos dos genomas de organismos resultaram em melhor compreensão das relações taxonômicas e das mudanças responsáveis pela evolução de espécies a partir de ancestrais comuns.

Agora que sabemos as sequências completas de nucleotídeos dos genomas de mais de 1.000 vírus, mais de 200 *archaea* e eubactérias e 28 eucariontes – mais de 44 bilhões de pares de nucleotídeos de DNA no total –, como podemos usar esta informação? Por si só, a sequência de uma molécula de DNA é totalmente não informativa. Em genética, usamos a dissecação mutacional para determinar as funções de várias unidades de DNA. Mas como podemos extrair informações das grandes quantidades de informações de sequências nos bancos de dados? Nesta seção, discutiremos resumidamente algumas das ferramentas que são usadas para “garimpar” informações de sequências de DNA, as atualmente disponíveis e um grande número de novas sequências que se acumulam diariamente.

Temos promessas de que a sequência do genoma humano resultará em novos enfoques na prática da medicina e que tratamentos serão ajustados ao genótipo de cada pessoa. Mulheres com uma mutação no gene *BRCA1* (câncer de mama 1) ou no gene *BRCA2*, por exemplo, provavelmente serão monitoradas por seus médicos quanto a câncer de mama de modo mais vigilante que outras mulheres (veja Cap. 24). Entretanto, na maioria, o tratamento médico baseado no genótipo da pessoa, à exceção de distúrbios hereditários clássicos, ainda é muito futurista. Assim, quais foram as contribuições imediatas da era da genômica? A resposta a esta pergunta parece clara – um conhecimento aumentado das relações evolutivas entre espécies e outros grupos taxonômicos. Comparando as sequências de nucleotídeos dos genomas de organismos, uma subdisciplina chamada **genômica comparativa**, cientistas documentaram muitas das mudanças responsáveis pela divergência de espécies de ancestrais comuns. A genômica comparativa é uma poderosa ferramenta nova para estudos de evolução. “Árvores” evolutivas – filogenias

que mostram as relações entre espécies e outros grupos taxonômicos – podem ser construídas a partir de sequências de DNA (veja Filogenias Moleculares, no Cap. 27). Examinaremos resumidamente algumas das mudanças que ocorreram nos genomas dos cereais e em alguns mamíferos nas duas últimas seções deste capítulo. O Cap. 27 apresenta uma discussão mais detalhada dos processos evolutivos.

Bioinformática

O conhecimento das sequências de nucleotídeos de genomas inteiros proporcionou muitas informações e um novo desafio. Como podemos extrair informações destas sequências? Este desafio faz parte de uma nova disciplina científica chamada **bioinformática** (biologia + informática). Sabemos que biologia é o estudo da vida. Informática é a ciência de reunir, manipular, estocar, recuperar e classificar informações registradas. Bioinformática envolve fazer todas estas coisas com a informação biológica, mais notadamente sequências de DNA e proteínas. Uma discussão compreensiva de bioinformática está além do escopo deste texto. Entretanto, vamos examinar resumidamente algumas das ferramentas usadas para estudar sequências de nucleotídeos do DNA e de aminoácidos em proteínas.

Suponha que você sequenciou um fragmento de restrição de DNA isolado de seu organismo favorito. Como você começaria a analisar a função desta molécula de DNA? Primeiro, você pode indagar se alguém já sequenciou esse DNA ou moléculas similares de DNA. Para responder a esta pergunta, você precisa de um programa de computação que possa pesquisar grandes bancos de DNA quanto a sequências similares. Programas de computação destinados a pesquisar sequências de bancos de dados foram primeiramente desenvolvidos nos anos 1980, e hoje existem programas destinados a quase tudo que você possa imaginar. Alguns dos programas mais populares foram desenvolvidos pelo Genetics Computer Group (GCG) na University of Wisconsin. Usaremos um par de seus programas para ilustrar de que modo sequências de nucleotídeos podem ser estudadas.

No início deste capítulo (veja Enfoque Técnico: GenBank), discutimos o uso do *web site* Entrez (<http://www.ncbi.nlm.nih.gov/entrez>) para pesquisar sequências de DNA similares a uma determinada sequência usando o programa megaBLAST. O outro programa que é usado para pesquisar rapidamente imensos bancos de dados é chamado de FASTA. Na pesquisa que fizemos no exercício Enfoque Técnico: GenBank, descobrimos que nossa sequência codificava parte da $\beta 9$ -tubulina de *Arabidopsis* e que ela era muito relacionada ao gene da $\beta 8$ -tubulina desta planta. A ferramenta BLAST mostrou-nos o quanto a sequência pesquisada era similar a cada um dos dois genes no GenBank (veja Enfoque Técnico, Fig. 2c).

Agora, vamos supor que nossa nova sequência não está representada no GenBank e vamos perguntar como podemos começar a analisá-la. A etapa mais elementar em ten-

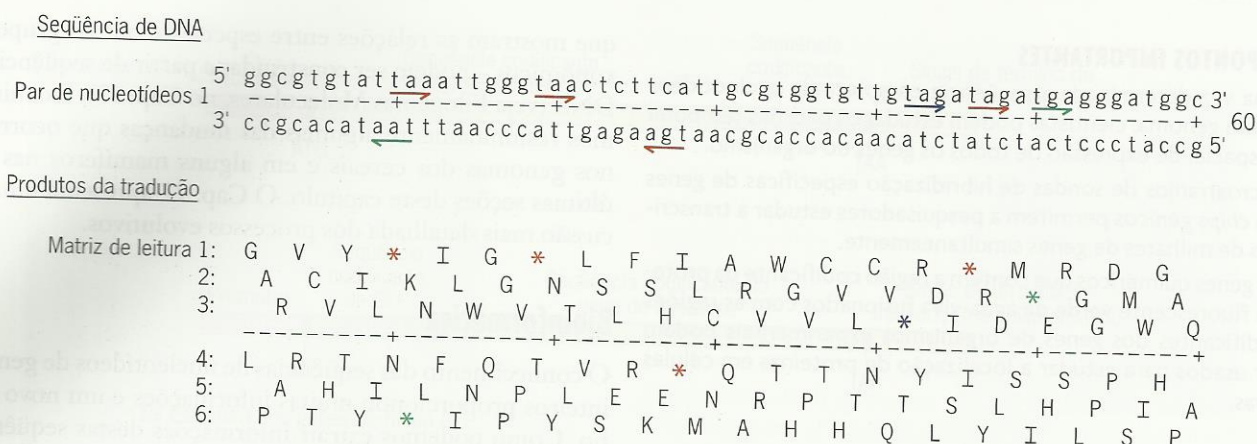


Fig. 16.21 ■ Ilustração do uso do programa Wisconsin GCG "Map" para identificar ORF traduzindo um segmento de 60 pares de bases de DNA de *C. reinhardtii* em todas as seis matrizes de leitura. Note a presença de códons de término de tradução em todas as matrizes de leitura exceto a número 5. Assim, é a única matriz de leitura que pode ser parte de uma ORF maior. Lembre que a tradução é sempre 5' → 3'; logo, os terminais amino estão à esquerda para os produtos de tradução 1–3 e à direita para os produtos 4–6. Os sítios de término de tradução são designados por asteriscos. Eles correspondem a trinças de término mostradas sublinhadas em vermelho, verde ou azul, dependendo da matriz de leitura.

tar identificar genes dentro de seqüências de nucleotídeos é procurar matrizes abertas de leitura (ORF) – seqüências que podem ser traduzidas em seqüências de aminoácidos sem encontrar nenhum dos códons de término de tradução na matriz de leitura. "Map" é um programa do GCG que pode ser usado para traduzir DNA bifilamentar em todas as seis matrizes de leitura (três em cada filamento de DNA). Usamos o programa Map para procurar ORF em um curto segmento de DNA da alga verde *Chlamydomonas reinhardtii* (Fig. 16.21). A abreviação padrão de aminoácidos com três letras (veja Quadro 12.1) não é muito boa quando se lida com grandes bancos de dados de proteína; portanto, o código de uma letra (veja Fig. 12.1) é preferido em análises de bioinformática.

Quando o curto segmento de DNA de *Chlamydomonas* é traduzido em todas as três matrizes de leitura, apenas a matriz de leitura 5 não tem códons de término (Fig. 16.21). Assim, qualquer ORF que inclua este segmento de DNA teria que estar na matriz de leitura 5. Logicamente, se estamos procurando genes, estaremos procurando ORF que são muito maiores do que a curta seqüência de DNA da Fig. 16.21. O programa desenvolvido para pesquisar seqüências de nucleotídeos de genes também pode triar promotores, sítios de ligação a ribossomos e outras seqüências conservadas. A presença de íntrons em genes eucarióticos torna as pesquisas de genes mais difíceis do que em procariontes. Pesquisas de genes eucarióticos envolvem encontrar sítios de corte de íntrons além de ORF e outras seqüências regulatórias. Hoje em dia, os programas de pesquisa de genes são ajustados a espécies individuais por características de seus genomas – por exemplo, composição de bases, uso de códons e preferências por determinadas seqüências em elementos regulatórios. O único modo de estar absolutamente certo de que um gene previsto é real e que seus íntrons foram identificados corretamente é isolar e seqüenciar todo um clone

de cDNA e comparar sua seqüência com a seqüência de um clone genômico. Atualmente existem dúzias de programas de previsão de genes. Eles têm nomes tais como GRAIL (Gene Recognition and Analysis Internet Link), GeneMark (um dos primeiros usados para pesquisar genes em *E. coli*), GeneScan, GeneFinder etc.

Uma característica comum de genomas de eucariontes é a presença de famílias de genes – conjuntos de genes que codificam proteínas muito similares (geralmente chamadas de isoformas). Essas proteínas geralmente têm funções redundantes ou superpostas. Para comparar todos os genes em uma família de genes, especialistas em bioinformática desenvolveram programas que irão alinhar várias seqüências de nucleotídeos ou aminoácidos, permitindo uma comparação visual direta de todos os membros de uma família de genes ou proteínas. Múltiplos alinhamentos em geral são especialmente úteis na identificação de seqüências conservadas de DNA que são elementos regulatórios importantes, tais como sítios de ligação de proteínas. Eles também são úteis na identificação de domínios funcionais importantes, e portanto conservados, dentro de proteínas. A Fig. 16.22 mostra o alinhamento de regiões aminotermiais das oito β -tubulinas de *Arabidopsis thaliana*. Note que dentro dos 60 a 61 aminoácidos mostrados existem cinco regiões de quatro ou mais aminoácidos que são conservados em todas as oito proteínas.

Genes com seqüências de nucleotídeos muito similares, tais como os nove genes codificantes das β -tubulinas mostrados na Fig. 16.22, geralmente, mas nem sempre, devem sua similaridade à sua evolução a partir de um gene ancestral comum. Tais genes são ditos **homólogos**. Note que "similar" e "homólogo" não são sinônimos. Os nove genes da β -tubulina de *Arabidopsis* são **homólogos**. Dois destes genes foram produzidos por um evento recente (na escala de tempo evolutivo) de duplicação gênica. Eles diferem em apenas 30

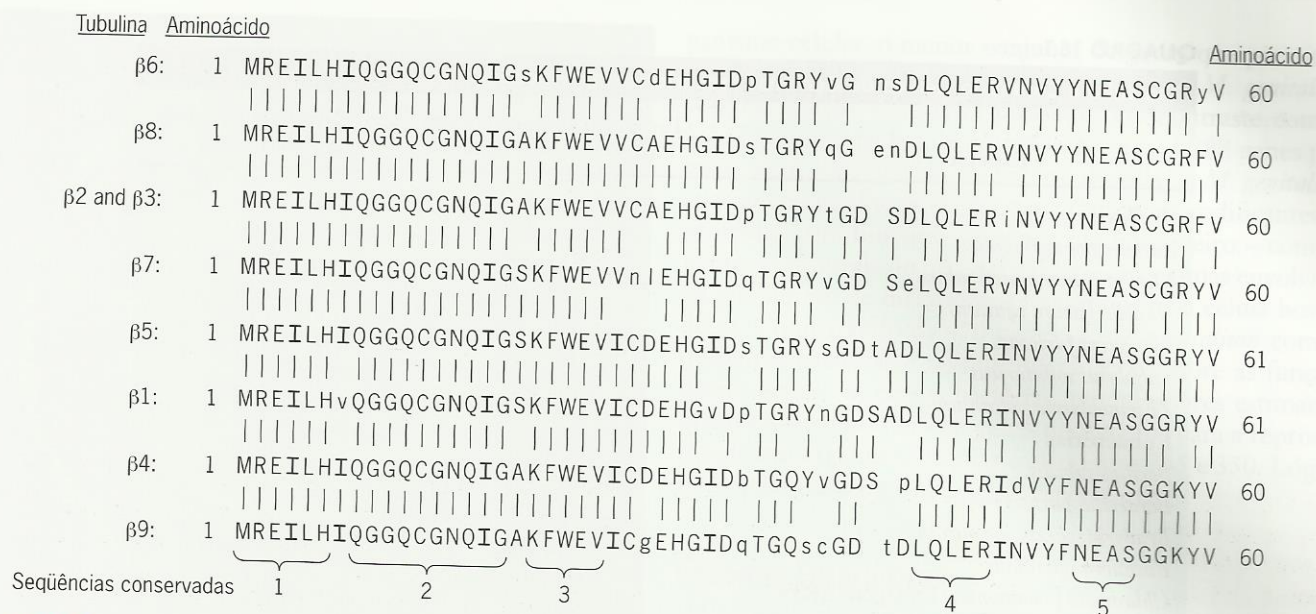


Fig. 16.22 ■ Ilustração de um alinhamento multissequência de proteína usando o código de uma letra para aminoácido (veja Fig. 12.1). O alinhamento, gerado com o programa GCG "PileUp", compara as regiões aminoterminais das oito β-tubulinas de *Arabidopsis*, que são codificadas por nove genes. Os genes *TUB2* (Tubulina Beta número 2) e *TUB3* codificam produtos idênticos (as tubulinas β2 e β3); estes dois genes diferem apenas nas posições correspondentes aos códons redundantes nas terceiras bases. As sequências mais similares são agrupadas. Os códigos de uma letra para aminoácidos estão em letras maiúsculas quando as sequências adjacentes são idênticas e em letras minúsculas quando as sequências adjacentes são diferentes.

pares de bases (correspondendo a bases redundantes em códons) e codificam o mesmo polipeptídeo. Tais genes também são denominados **parálogos** – genes homólogos dentro de uma espécie. Genes homólogos presentes em espécies diferentes são chamados de **ortólogos**. Os genes de tubulina de *Arabidopsis* e *Chlamydomonas* são ortólogos.

Genomas Procarióticos

Em 1995, *Haemophilus influenzae* tornou-se o primeiro organismo celular a ter seu genoma sequenciado em sua totalidade. Em janeiro de 2005, as sequências completas dos genomas de 20 *archaea* e 190 bactérias estavam disponíveis em bancos de dados públicos, e este número será muito maior quando este livro estiver sendo impresso. Os genomas variam de tamanho: 490.885 pb para *Nanoarchaeum equitans*, um simbionte obrigatório; 580.070 pb para *Mycoplasma genitalium*, que se acredita ter o menor genoma de qualquer bactéria não simbiótica; 4.411.529 pb para *Mycobacterium tuberculosis*, a causa de mais mortes humanas que qualquer outra bactéria infecciosa; 4.639.221 pb para *Escherichia coli*, o microrganismo celular mais bem conhecido; 9.105.828 pb para *Bradyrhizobium japonicum*, uma bactéria do solo capaz de colonizar nódulos de raízes de plantas. O tamanho e o conteúdo gênico previsto de alguns genomas procarióticos são mostrados no Quadro 16.1.

Antes de discutir genomas bacterianos em maiores detalhes, precisamos destacar que os tamanhos dos genomas

são altamente variáveis mesmo dentro de espécies, variando de uma linhagem para outra. De fato, estudos em diferentes isolados de *E. coli*, *Prochlorococcus marinus* e *Streptococcus coelicolor* documentaram variações no tamanho do genoma de até um milhão de pares de nucleotídeos entre linhagens diferentes da mesma espécie.

Os genomas de *archaea* são de interesse particular em razão da capacidade destes organismos de sobreviver em ambientes extremamente hostis. Seus genomas devem ajudar-nos a determinar como eles se adaptaram a tais condições extremas. A espécie *Sulfolobus*, por exemplo, pode crescer em temperaturas de até 85°C em solo muito ácido (pH = 1). Algumas espécies de *Geobacter* podem crescer em temperaturas de até 121°C, bem acima do ponto de fervura da água, o que lhes permite viver em locais hidrotérmicos no fundo dos oceanos. Outras *archaea*, as halófilas, adaptaram-se a ambientes muito salinos tais como o Great Salt Lake em Utah. Determinando como os genomas destes organismos mudaram para lhes permitir sobreviver em ambientes tão extremos, desenvolveremos melhor compreensão do processo de evolução e da vida em nosso planeta.

Dos genomas bacterianos sequenciados até hoje, a sequência do genoma de *E. coli* (Fig. 16.23) sem dúvida foi o que causou o maior entusiasmo entre os geneticistas. *E. coli* é o organismo celular mais estudado e mais bem compreendido de nosso planeta. Geneticistas, bioquímicos e biólogos moleculares usaram *E. coli* como o organismo modelo preferido por décadas. A maior parte do que sabemos sobre

QUADRO 16.1

Tamanho e Conteúdo Gênico de Genomas Procarióticos Seleccionados		
Espécie	Tamanho Genômico em Pares de Nucleotídeos	Número Previsto de Genes
Archaea		
<i>Archaeoglobus fulgidus</i>	2.178.400	2.407
<i>Methanosarcina acetivorans</i>	5.752.492	4.540
<i>Nanoarchaeum equitans</i>	490.885	563
<i>Pyrococcus furiosus</i>	1.908.256	2.065
<i>Sulfolobus solfataricus</i>	2.992.245	2.977
<i>Thermoplasma volcanium</i>	1.584.804	1.449
Eubactérias		
<i>Bacillus subtilis</i>	4.214.630	4.106
<i>Bordetella parapertussis</i>	4.773.551	4.402
<i>Bradyrhizobium japonicum</i>	9.105.828	8.317
<i>Buchnera aphidicola</i>	615.590	504
<i>Chlamydia pneumoniae</i> , linhagem AR39	1.229.853	1.110
<i>Escherichia coli</i> , linhagem K12	4.639.221	4.288
<i>Escherichia coli</i> , linhagem O157	5.528.970	5.349
<i>Haemophilus influenzae</i>	1.830.138	1.709
<i>Legionella pneumophila</i> , linhagem Paris	3.503.610	3.082
<i>Mycobacterium tuberculosis</i> , linhagem CDC	4.403.836	4.187
<i>Mycobacterium genitalium</i>	580.074	480
<i>Neisseria meningitidis</i>	2.184.406	2.121
<i>Phytoplasma asteris</i>	860.631	754
<i>Pseudomonas syringae</i>	6.397.126	5.471
<i>Rickettsia typhi</i>	1.111.496	838
<i>Salmonella typhimurium</i>	4.857.432	4.452
<i>Staphylococcus aureus</i> , linhagem MW2	2.820.462	2.632
<i>Streptomyces coelicolor</i>	8.667.507	7.826
<i>Ureaplasma urealyticum</i>	751.719	611
<i>Yersinia pestis</i> , linhagem KM	4.600.755	4.086

Dados do Center for Biological Sequence Analysis Genome Atlas Database em <http://www.cbs.dtu.dk/services/GenomeAtlas> de 14 de janeiro de 2005.

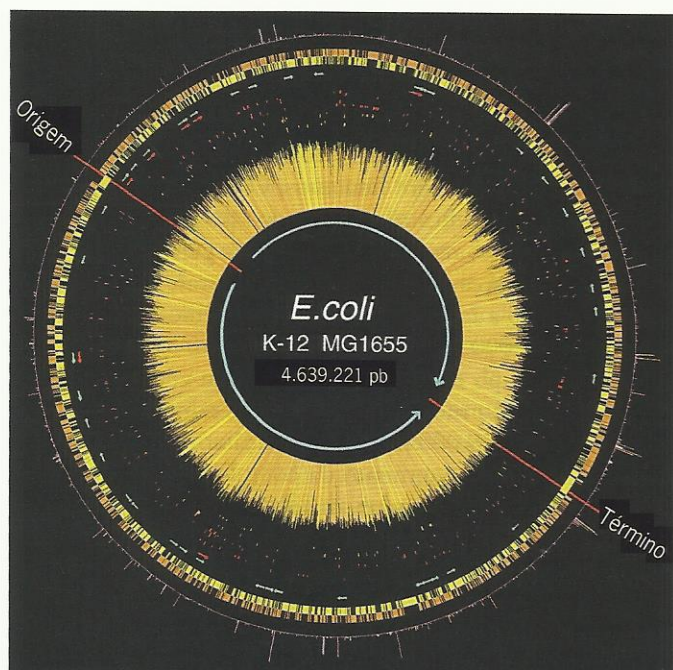
genética de bactérias foi aprendida de pesquisas com *E. coli*. Assim, a publicação da sequência completa do genoma de *E. coli* em 1997 foi um marco significativo na história da genética.

O genoma de *E. coli* contém 4.288 supostos genes codificantes de proteínas. Desses genes previstos, cerca de um terço são genes bem estudados que codificam produtos conhecidos, enquanto 38% têm função desconhecida. A Fig. 16.24 mostra uma classificação dos genes de acordo com a função do produto gênico. A distância média entre genes (tamanho das regiões intergênicas) no genoma de *E. coli* é de 118 pb. Genes conhecidos e supostos que especificam proteínas e RNA estáveis constituem 87,8% e 0,8% do genoma, respectivamente, e elementos repetitivos não codificantes correspondem a 0,7% do genoma. Assim, 10,7% do genoma deve envolver sequências regulatórias, sequências com outras funções desconhecidas e talvez algumas sequências de nucleotídeos não funcionais.

Outros genomas bacterianos exibem a mesma densidade de genes que *E. coli*, cerca de um gene por kb de DNA. Isto contrasta muito com a variação na densidade de genes observada nos genomas de eucariontes, em que genomas maiores têm densidades gênicas muito baixas.

Os genomas de *M. tuberculosis* (tuberculose), *Mycobacterium leprae* (hanseníase), *Corynebacterium diphtheriae* (difteria), *Legionella pneumophila* (doença dos legionários), *Bacillus anthracis* (antraz), *Yersinia pestis* (peste bubônica) e outras bactérias infecciosas também são de grande interesse em virtude da patogenicidade destes organismos e da esperança de que uma compreensão completa de seu metabolismo irá sugerir meios para evitar estas doenças geralmente fatais. A necessidade de novos meios para combater patógenos bacterianos foi acentuada pela recente evolução de linhagens destas bactérias infecciosas resistentes a antibióticos (Cap. 8).

O genoma de *M. genitalium* é de especial interesse porque pode aproximar-se do “conjunto gênico mínimo” de um or-



Chave: — genes de tRNA; — genes de rRNA;
— origem e término de replicação

Fig. 16.23 ■ Mapa de sequência de bases do cromossomo de *Escherichia coli*. As setas em azul marcam as metades do cromossomo atravessadas pelas forquilha de replicação. O círculo concêntrico externo dá as posições dos genes codificantes de proteínas similares às proteínas do bacteriófago. O segundo círculo concêntrico mostra a localização de genes que são transcritos no sentido horário (dourado) de um filamento ou no sentido anti-horário (amarelo) de um filamento complementar. O sol no centro é um histograma no qual o tamanho de cada raio é proporcional a aleatoriedade de uso do códon dentro de cada sequência codificante.

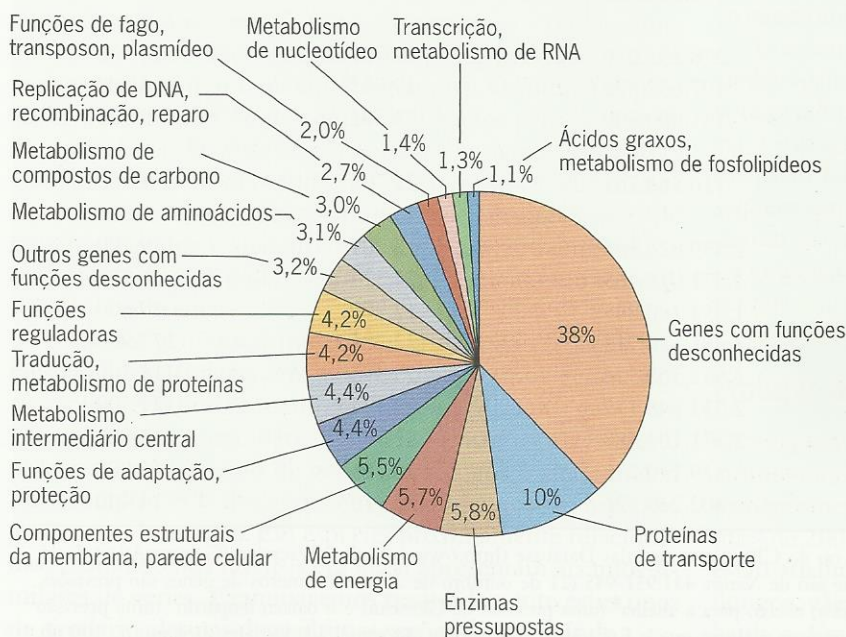


Fig. 16.24 ■ Classificação dos 4.288 genes conhecidos e supostos de *E. coli* de acordo com a função. A classe maior de genes conhecidos codifica proteínas envolvidas em transporte. Entretanto, note que 38% dos genes codificam proteínas de função desconhecida.

ganismo celular, o menor conjunto de genes que permitiria a reprodução de uma célula. O genoma de *M. genitalium* contém apenas 470 genes previstos, em contraste com os 1.727 genes previstos de *H. influenzae* e os 4.288 genes previstos de *E. coli*. O ciclo parasitário de vida de *M. genitalium* aparentemente lhe permitiu descartar genes codificantes de proteínas com funções dadas por seu hospedeiro – componentes estruturais de paredes celulares e enzimas envolvidas na biossíntese de metabólitos fornecidos pela célula hospedeira. Comparando os 470 genes de *M. genitalium* com os de outras bactérias e usando informações sobre as funções desses genes em outras bactérias, pesquisadores estimaram que o número mínimo de genes necessários para a reprodução de um organismo celular é algo entre 265 e 350. Lógico que esta estimativa é altamente especulativa.

Genomas Eucarióticos

A levedura do pão, *Saccharomyces cerevisiae*, foi o primeiro organismo eucariótico a ter todo o seu genoma sequenciado. A sequência completa de 12.068 kb do genoma de *S. cerevisiae* foi montada em 1996 por uma colaboração internacional de cerca de 600 cientistas trabalhando na Europa, na América do Norte e no Japão. O genoma da levedura contém 5.885 potenciais genes codificantes de proteínas, cerca de 140 genes especificando RNA ribossômicos, 40 genes para moléculas de pequeno RNA nuclear e mais de 200 genes para tRNA. Pesquisadores sistematicamente geraram deleções de essencialmente todos (5.916, ou 96,5%) dos 6.268 genes previstos no genoma da levedura. Dos genes testados, 1.105 (18,7%) foram considerados essenciais para crescimento em meio rico em glicose, isto é, as deleções nestes genes eram letais. Algumas deleções não causam letalidade porque o ge-

noma da levedura contém muitos genes duplicados. Ambas as cópias destes genes têm que ser deletadas para haver um efeito letal. Muitos outros genes de levedura podem ser deletados sem matar o organismo. Entretanto, nocautes desses genes estão geralmente associados a mudanças na morfologia ou crescimento comprometido.

As seqüências dos genomas de outros sistemas modelos eucarióticos vieram logo em seguida. A seqüência de 99% do genoma do verme *Caenorhabditis elegans* foi publicada em 1998, e as seqüências quase completas dos genomas da mosca-das-frutas *Drosophila melanogaster* e da planta modelo *Arabidopsis thaliana* foram publicadas em 2000. A liberação dos dois primeiros rascunhos da seqüência do genoma humano em 2001 provavelmente recebeu maior cobertura dos noticiários internacionais do que qualquer outro evento na história da biologia. A publicação das seqüências dos genomas de vários outros eucariontes logo se seguiu, e estão em andamento projetos de seqüenciamento de muitas outras espécies. Como já foi mencionado, uma seqüência quase completa do genoma humano foi publicada em 2004.

O que aprendemos com todas estas seqüências? Em contraste com os genomas de *archaea* e eubactérias, a densidade de genes varia muito entre espécies eucarióticas diferentes, indo de um gene por 1.900 pb na levedura do pão a um gene por 127.900 pb (145.000 pb se for incluída a heterocromatina não seqüenciada) em humanos. O tamanho do genoma e o conteúdo de genes são mostrados para os genomas de alguns eucariontes no Quadro 16.2. Os genomas de eucariontes unicelulares são como o de leveduras, com um gene para cada 1.000 a 2.000 pb. A densidade gênica diminui para um gene por 4.000 a 5.000 pb em *Arabidopsis* e *C. elegans*, um gene por 9.500 pb em *D. melanogaster*, um gene por 15.000 pb no baiacu e é a menor em mamíferos com um gene para cada 115.000 a 129.000 pb. A diminuição observada na densidade gênica com o aumento da complexidade relacionada ao desenvolvimento levanta dúvidas sobre as funções do DNA não codificante. Como já foi mencionado (veja Fig. 16.16), recentemente foi organizado um *ENCyclopedia Of DNA Elements* (ENCODE) Project Consortium internacional para investigar este material.

QUADRO 16.2

Tamanho e Conteúdo Gênico Previsto de Genomas Eucarióticos Selecionados

Espécie	Nome Comum	Tamanho do Genoma em Pares de Nucleotídeos	Número Previsto de Genes*	Densidade Gênica (pb/gene)††
Protistas				
<i>Encephalitozoon cuniculi</i>	microsporídeo	2.497.519	1.996	1.300
<i>Plasmodium falciparum</i>	protozoário da malária	22.820.308	5.317	4.300
Fungos				
<i>Candida glabrata</i>	leveduras (infeciosas)	12.280.356	5.272	2.300
<i>Saccharomyces cerevisiae</i>	levedura do pão	12.057.909	6.268	1.900
Nematódeos				
<i>Caenorhabditis elegans</i>	nematelmintos	100.291.841	20.516	4.900
Insetos				
<i>Anopheles gambiae</i>	mosquito	278.253.050	14.707	18.900
<i>Apis mellifera</i>	abelha	197.657.892	29.832 [#]	6.600
<i>Drosophila melanogaster</i>	mosca-das-frutas	131.000.899	13.792	9.500
Plantas				
<i>Arabidopsis thaliana</i>	agrião	116.566.763	25.706	4.500
Vertebrados				
<i>Canis familiaris</i>	cão	2.359.826.366	18.201	129.700
<i>Danio rerio</i>	peixe zebra	1.571.018.465	23.524	66.800
<i>Gallus gallus</i>	galinha	1.054.180.845	17.709	59.600
<i>Homo sapiens</i>	homem	2.851.330.913	22.287	127.900
<i>Mus musculus</i>	camundongo	2.932.368.526	25.396	115.500
<i>Pan troglodytes</i>	chimpanzé	2.733.948.177	22.524	121.500
<i>Rattus norvegicus</i>	rato norueguês	2.571.104.688	22.159	115.800
<i>Takifugu rubripes</i>	baiacu japonês	329.140.338	20.796	15.800
<i>Tetraodon nigroviridis</i>	baiacu leopardo	402.240.326	27.918 ^o	14.400

Os dados são do *web site* Ensembl (<http://www.ensembl.org>) ou do CBS Genome Atlas Database (<http://www.cbs.dtu.dk/services/GenomeAtlas>) de 15 de janeiro de 2005, exceto os dados do genoma humano, que são de *Nature* 431:931-945 (21 de outubro de 2004). *Números de genes são previsões do Ensembl (menos pseudogenes quando são disponíveis dados) exceto para a abelha[#] (uma previsão de Genscan) e o baiacu leopardo^o (uma previsão de Genoscope). Note que os últimos programas deram estimativas maiores que o Ensembl. ††Os valores são arredondados para próximo de 100 pb.

Um motivo para a diminuída densidade gênica nos genomas eucarióticos maiores é que esses genomas contêm quantidades consideráveis de DNA repetitivo (veja Cap. 9). A levedura do pão contém muito pouco DNA repetitivo, embora cerca de 30% de seus genes estejam duplicados. Em contraste, os genomas de eucariontes multicelulares contêm muito DNA repetitivo, e a quantidade deste material está, na maioria dos casos, diretamente relacionada com o tamanho do genoma. Por exemplo, apenas cerca de 10% do pequeno genoma de *C. elegans* consiste em DNA moderadamente repetitivo, enquanto cerca de 45% dos grandes genomas de mamíferos são compostos de seqüências de DNA moderadamente repetidas. A maioria destas seqüências moderadamente repetitivas é derivada de elementos genéticos de transposição (veja Cap. 18).

DNA altamente repetitivo também é mais abundante nos genomas maiores. Entretanto, não há correlação direta entre a quantidade de DNA altamente repetitivo e o tamanho do genoma. Às vezes, espécies proximamente relacionadas diferem de maneira significativa na quantidade de DNA altamente repetitivo. Por exemplo, 18% do genoma de *D. melanogaster* consiste em DNA altamente repetitivo, enquanto 45% do DNA em *D. virilis* é altamente repetitivo. Grande parte do DNA altamente repetitivo na maioria das espécies, incluindo humanos, está presente nas regiões de cromossomos flanqueadoras dos centrômeros (heterocromatina centromérica) e nos telômeros. Esse DNA é difícil de seqüenciar. De fato, a maioria do DNA não seqüenciado no genoma humano, 472 milhões de pares de bases, consiste em seqüências altamente repetitivas, e 24 dos espaços na seqüência do genoma humano correspondem a blocos de heterocromatina centromérica nos 24 cromossomos.

Íntrons são um componente significativo do DNA eucariótico e são mais prevalentes e maiores nos grandes genomas eucarióticos. As regiões intergênicas também são maiores nos genomas eucarióticos grandes. Em contraste, o número de domínios distintos de proteínas – regiões funcionais de proteínas – codificadas por genes não parece variar muito com o tamanho do genoma. Os números previstos de domínios de proteínas codificados pelos genomas de *A. thaliana*, *D. melanogaster* e seres humanos são 1.012, 1.035 e 1.262, respectivamente. Entretanto, seres humanos e outros vertebrados usam mais as vias alternativas de recomposição de transcrito (veja Cap. 21) para embaralhar esses domínios em mais combinações, aumentando a diversidade de proteínas.

Pesquisas genômicas mostraram que mesmo espécies distantemente correlatas têm muitos dos mesmos genes. Por exemplo, 18% dos genes de *Arabidopsis* e 50% dos genes em *Drosophila* têm homólogos humanos. Entre espécies muito próximas, a proporção de genes homólogos é ainda maior. Por exemplo, 99% dos genes em camundongos são homólogos a genes humanos. Em espécies proximamente relacionadas, cromossomos inteiros em geral apresentam arranjos similares de genes. Examinaremos resumidamente estes tipos de similaridades nas duas últimas seções deste capítulo.

Evolução do Genoma em Cereais

Os cultivos de cereais fornecem grande parte dos alimentos para seres humanos e seus animais domésticos, os quais fornecem alimentos humanos adicionais. Assim, a produtividade acentuada de cereais é um componente importante do esforço para alimentar a população humana em constante expansão em nosso planeta. O aumento do conhecimento sobre os genomas destas espécies agrônômicas é a chave para se atingir este objetivo. Análises comparativas recentes dos genomas de várias espécies de cereais indicam que grande parte das informações obtidas por mapeamento e seqüenciamento do menor desses genomas, o genoma de 400 mb de arroz, será diretamente aplicável a outros cereais devido à conservação da estrutura genômica dessas espécies.

Quando Graham Moore e colaboradores compararam os mapas genéticos de alta densidade dos cromossomos de vários cereais, eles descobriram que, a despeito de grandes diferenças no tamanho do genoma e no número de cromossomos, as relações de ligação de blocos de seqüências únicas de DNA e genes conhecidos são marcadamente conservadas. Em contraste, as quantidades e os locais das seqüências repetidas de DNA são altamente variáveis.

A marcante conservação da estrutura genômica nos cereais é mais claramente ilustrada desenhando-se o genoma do arroz em forma circular e alinhando os blocos de genes conservados na outra espécie com o genoma do arroz (Fig. 16.25). A disposição circular dos genomas de cereais não implica nenhuma circularidade de cromossomos ancestrais; ela apenas permite o alinhamento máximo de blocos homólogos de genes. O alinhamento também destaca a presença de cópias duplicadas de cada bloco de genes no genoma do milho, indicando que o milho evoluiu de um ancestral tetraplóide. Curiosamente, um conjunto de genes está amplamente presente nos pequenos cromossomos do milho, e o segundo grupo está presente primariamente nos grandes cromossomos.

As estruturas conservadas dos genomas de cereais devem ajudar agricultores em suas tentativas de produzir variedades com produtividade aumentada, resistência a pestes, tolerância a estiagem e outras características desejadas. Devido à estrutura conservada do genoma, as informações obtidas do seqüenciamento do genoma relativamente pequeno do arroz devem ser mais facilmente aplicadas a projetos de engenharia genética e cruzamentos em outros cultivos de espécies de cereais.

Evolução do Genoma em Mamíferos

Os genomas de mamíferos exibem conservação da estrutura cromossômica similar à observada em gramíneas. Embora o mapeamento do material genético esteja sendo feito em mais de 200 espécies de mamíferos, mapas de alta densidade estão atualmente disponíveis apenas para humanos, camundongos, cães, ratos e alguns animais tais como porcos e rebanhos. Esses mapas cromossômicos detalhados podem ser

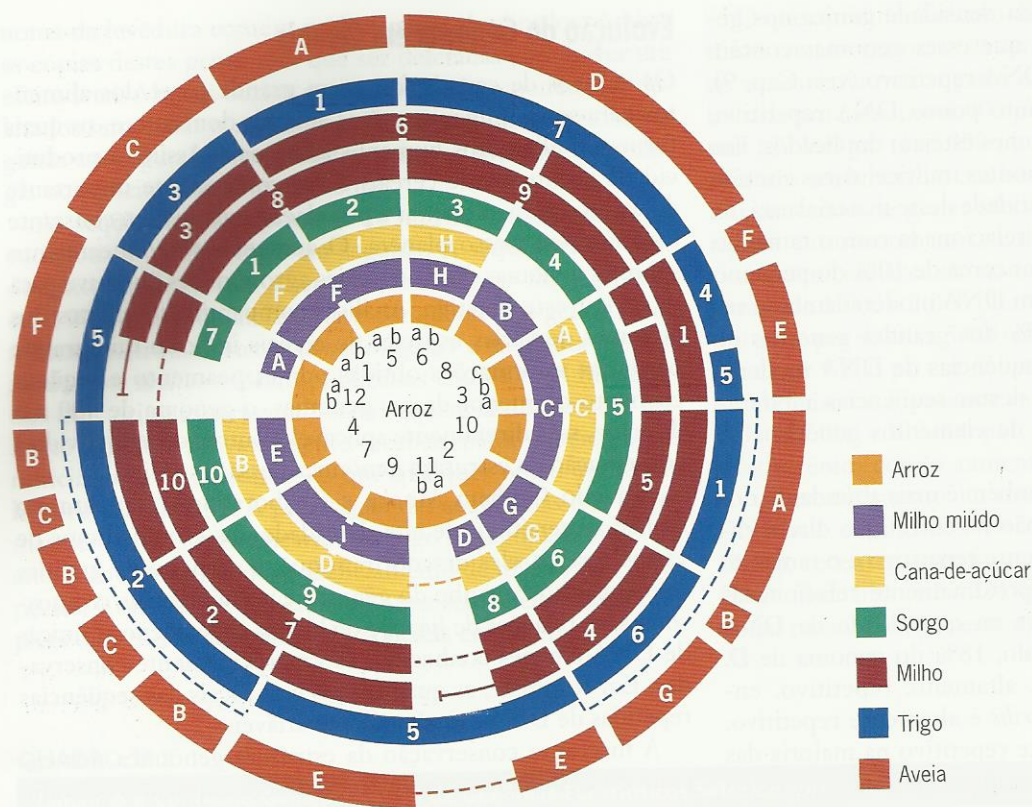


Fig. 16.25 ■ Mapa comparativo simplificado dos genomas de sete cereais. Os cromossomos e segmentos de cromossomos (indicados por letras maiúsculas) dos vários cereais são alinhados com os cromossomos de arroz, a gramínea com o menor genoma (centro). O genoma de milho tem duas cópias similares de cada bloco de genes e assim ocupa dois anéis do círculo. As linhas externas tracejadas conectam segmentos adjacentes de cromossomos de trigo. Segmentos similares de cromossomos no genoma de aveia não estão conectados por linhas tracejadas por simplificação.

usados para demonstrar as relações conservadas de ligação de genes em espécies em que tais mapas estão disponíveis. Para outras espécies, um procedimento chamado **colorir cromossomos** está sendo usado para análise comparativa de genomas. Colorir cromossomos é uma variação da hibridização *in situ* com fluorescência (FISH; veja Enfoque Técnico: Hibridização *In Situ* no Cap. 9) na qual cromossomos são “coloridos” com cores diferentes usando-se sondas de hibridização de DNA marcadas com corantes fluorescentes que emitem luzes de comprimentos de onda diferentes.

Em estudos de genômica comparativa, seqüências de DNA de uma espécie são usadas para colorir os cromossomos de uma espécie correlata (Fig. 16.26). Tais experimentos de coloração de cromossomos interespecíficos são chamados de experimentos de Zoo-FISH. Eles geralmente são feitos em condições de baixa restrição de hibridização, o que permite a detecção de hibridização cruzada entre os filamentos parcialmente complementares de genes homólogos.

Alguns dos estudos mais interessantes de pintar cromossomos usaram as seqüências em uma biblioteca genômica específica de cromossomos (Cap. 15) de uma espécie para pintar os cromossomos de espécies correlatas. Corantes fluorescentes diferentes podem ser usados para marcar seqüências de dois ou mais cromossomos diferentes de uma espécie, e tais seqüências marcadas com fluorescência podem ser usadas para “pintar” os cromossomos de espécies correlatas. Bibliotecas específicas de cromossomos estão disponíveis para todos os 24 cromossomos humanos, e seqüências dessas

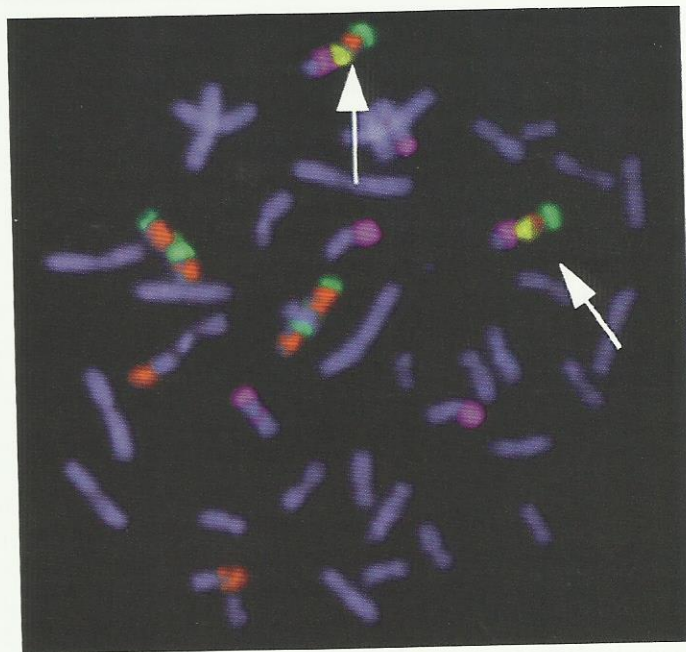


Fig. 16.26 ■ Ilustração do uso de colorir cromossomos (veja texto) para estudar correlações evolutivas entre cromossomos. A hibridização *in situ* fluorescente feita em cromossomos do gibão *Hylobates lar* usando-se sondas preparadas de bibliotecas de clones específicos de cromossomo dos cromossomos humanos 5 (laranja), 16 (verde), 17 (rosa) e 22 (amarelo). As setas indicam o cromossomo 8 de gibão, que contém seqüências que hibridizam com sondas fluorescentes derivadas de todos os quatro destes cromossomos humanos.

bibliotecas foram usadas para “pintar” os cromossomos de várias espécies correlatas, incluindo a maioria dos primatas e alguns mamíferos um pouco distantemente correlatos.

Reverendo a ligação comparativa e os dados de pintar cromossomos, Bhanu Chowdhary e colaboradores concluíram que a evolução de cromossomos de mamíferos envolve três classes de sintenia conservada. (Sintenia é a presença de genes no mesmo cromossomo.) As três classes são (1) conservação de cromossomos inteiros, (2) conservação de grandes segmentos de cromossomos e (3) a união de segmentos de cromossomos diferentes para produzir nova sintenia. Cada tipo de evolução cromossômica é ilustrado na Fig. 16.27.

A sintenia de genes no cromossomo humano 17 é conservada no cromossomo 12 do porco e no cromossomo 19 do gado (Fig. 16.27a). Padrões similares de conservação dos cromossomos humanos 13 e 20 foram observados nestas espécies.

O cromossomo humano 2 fornece um exemplo da segunda classe de sintenia conservada, a conservação de grandes segmentos de cromossomos (Fig. 16.27b). Partes importantes dos braços longo e curto do cromossomo humano 2 são conservadas em dois cromossomos do porco e do gado

(também em cavalo e gato, não mostrados). Os genes no cromossomo humano 2 são mais dispersos nos cromossomos do cervo indiano e do camundongo (não mostrados). Os cromossomos humanos 4, 5, 6, 9 e 11 também têm grandes segmentos cromossômicos conservados em outros mamíferos estudados.

O terceiro padrão de sintenia conservada, a união de segmentos de cromossomos para produzir nova sintenia, é ilustrado pelos cromossomos humanos 3 e 21 (Fig. 16.27c). O cromossomo 13 do porco parece conter os genes de ambos os cromossomos humanos dispostos como se tivessem simplesmente fundidos. No gado, um bloco de genes presentes no cromossomo humano 3 está presente em um cromossomo separado (22), enquanto a maioria do resto dos genes do cromossomo humano 3 estão em um cromossomo (1) com os genes do cromossomo humano 21. Outros exemplos deste padrão de sintenia conservada envolvem os cromossomos humanos 12 e 22, 14 e 15 e 16 e 19.

A despeito da ampla conservação de blocos de genes ilustrada pelos exemplos já discutidos, comparações mais detalhadas da estrutura cromossômica em mamíferos mostram claramente que vários rearranjos cromossômicos ocorreram

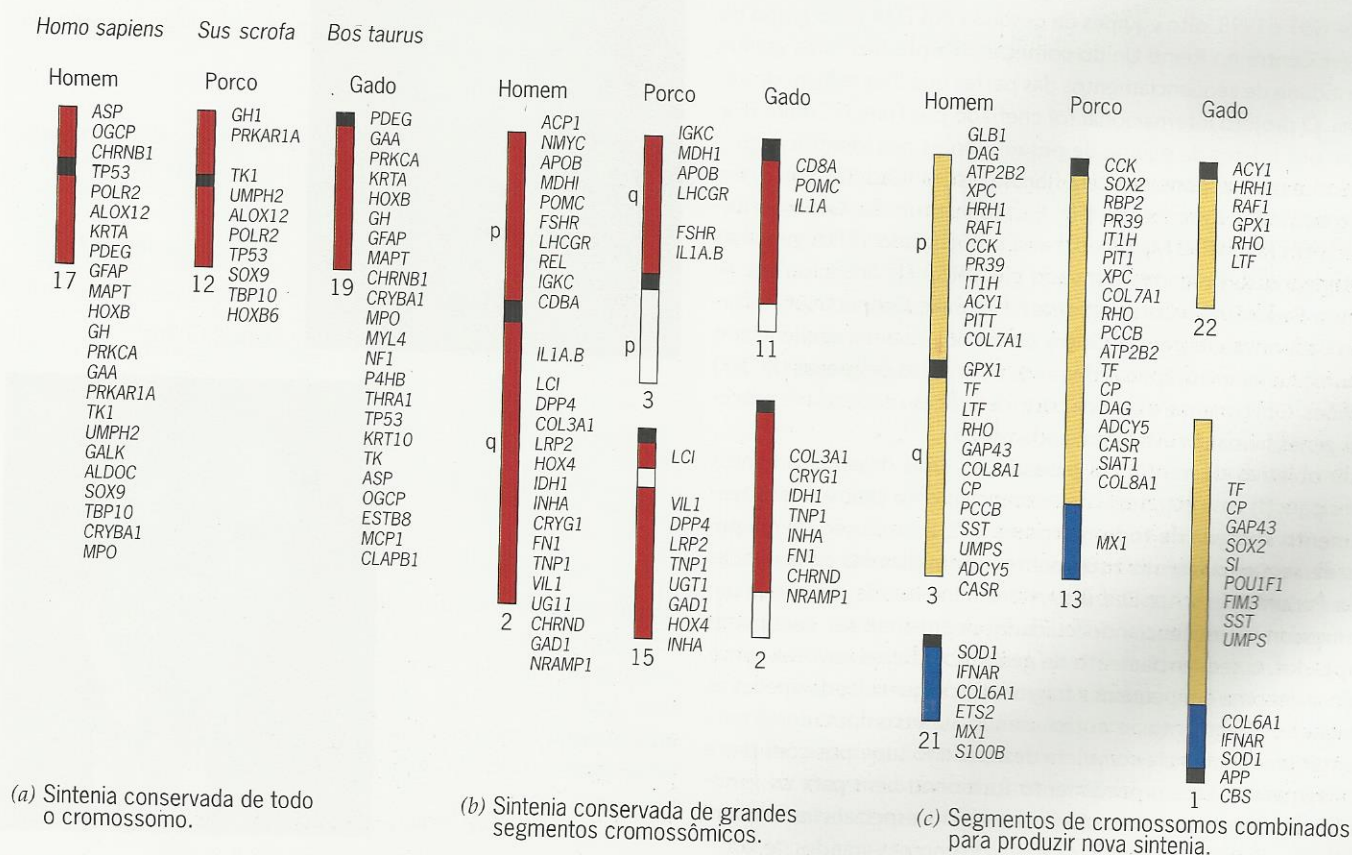


Fig. 16.27 ■ Evolução de cromossomos em mamíferos. São ilustrados exemplos das três classes de sintenia conservada (blocos conservados de genes ligados). (a) O cromossomo humano 17 é um exemplo da conservação de cromossomos inteiros. Note que mesmo com a sintenia conservada ocorrem inversões, mudando a ordem de alguns genes. (b) O cromossomo humano 2 dá um exemplo da conservação de grandes segmentos cromossômicos. (c) Os cromossomos humanos 3 e 21 ilustram a formação de nova sintenia por fusão de cromossomos. As localizações cromossômicas de alguns genes são mostradas à direita dos cromossomos. Menos genes foram mapeados em porcos do que no gado ou em humanos. Designações cromossômicas espécie-específicas são dadas abaixo dos cromossomos.

durante a evolução de espécies muito próximas. Uma comparação de cromossomos humanos com os cromossomos do gibbon de bochecha branca, *Hylobates concolor*, revelou que pelo menos 21 translocações ocorreram durante a evolução destas duas espécies de seus ancestrais comuns. Inversões e outros rearranjos intracromossômicos são especialmente prevalentes nos genomas de espécies proximamente correlatas.

PONTOS IMPORTANTES

- A genômica comparativa – comparar as seqüências de nucleotídeos dos genomas – forneceu novas informações sobre as relações entre vários grupos taxonômicos.
- Bioinformática é a ciência de estocar, comparar e extrair informações de sistemas biológicos, especialmente DNA e seqüências de proteínas.
- As seqüências de nucleotídeos de diversos organismos procarióticos deram *insights* quanto a sua adaptação a nichos ambientais únicos.
- À medida que organismos eucarióticos aumentaram sua complexidade, a proporção de seus genomas que codificam proteínas diminuiu, e a função da maior parte do DNA não codificante é desconhecida.
- A genômica comparativa revelou uma conservação marcante de sintenia em espécies eucarióticas correlatas, tais como mamíferos e cereais.



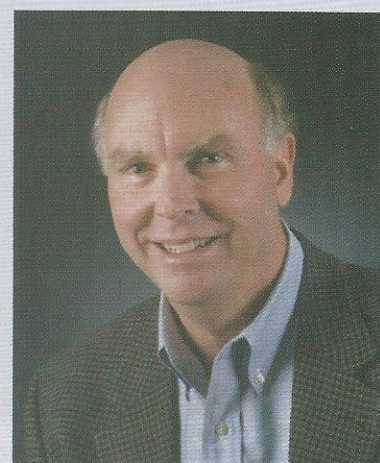
UM MARCO NA GENÉTICA: Dois Rascunhos da Seqüência do Genoma Humano

O Projeto do Genoma Humano foi iniciado em 1990, com a meta de seqüenciar todo o genoma humano – todos os 3 bilhões de pares de nucleotídeos – em 2005, a um custo estimado de US\$3 bilhões. Durante 1997 e 1998, oito equipes de pesquisa nos EUA e um grupo do Sanger Centre no Reino Unido começaram a produzir uma grande quantidade de seqüenciamentos das partes que lhes cabiam do genoma. O projeto internacional foi chefiado por Francis Collins (Fig. 1), um dos líderes da equipe de pesquisadores que identificaram o gene mutante responsável pela fibrose cística (Cap. 17). Então, em maio de 1998, J. Craig Venter (Fig. 1), cujo Institute for Genomic Research em Rockville, Maryland, havia seqüenciado vários genomas bacterianos, surpreendeu o mundo científico. Ele anunciou que se uniria a PerkinElmer Corporation de Norwalk, Connecticut, e formou uma nova empresa, a Celera Genomics, que iria seqüenciar o genoma humano em apenas três anos a um custo de apenas US\$300 milhões. Obviamente, a corrida para identificar, seqüenciar e patentear genes humanos estava em andamento.

O objetivo de Venter foi baseado em dois desenvolvimentos principais: (1) uma estratégia de seqüenciamento chamada **seqüenciamento shotgun de todo o genoma** e (2) a produção de **máquinas de seqüenciamento totalmente automatizadas** mais rápidas pela PerkinElmer's Applied Biosystems. Em vez de percorrer um cromossomo, seqüenciando cuidadosamente clones genômicos mapeados, o seqüenciamento de genomas inteiros envolve cortar todo o genoma em pequenos fragmentos, seqüenciando apenas as pontas dos fragmentos e então usando supercomputadores para montar uma seqüência completa descobrindo superposições entre os fragmentos. Este procedimento funcionou bem para os genomas bacterianos muito menores, mas alguns especialistas estavam céticos quanto à exequibilidade com os genomas grandes de eucariotes. Enquanto genomas bacterianos contêm pouco DNA repetitivo, genomas eucarióticos contêm grandes quantidades de DNA altamente repetitivo. Tais seqüências altamente repetitivas podem corresponder a muitos fragmentos e portanto obstruir o processo de montagem da seqüência pelo computador.



Francis Collins



J. Craig Venter

Fig. 1 ■ Os líderes das equipes pública e privada envolvidos no seqüenciamento do genoma humano: Francis Collins, diretor do Human Genome Project, e J. Craig Venter, presidente fundador da Celera Genomics, a empresa privada formada para seqüenciar o genoma humano.



UM MARCO NA GENÉTICA (continuação)

A máquina de seqüenciamento de DNA desenvolvida por PerkinElmer podia separar os produtos de 96 reações didesoxi de término de cadeia (Cap. 15) simultaneamente e podia operar direto por 24 horas. Todas as funções, incluindo carregar o gel, colocar amostra, eletroforese, coleta de dados e análise dos dados, eram totalmente automatizadas. Os seqüenciadores eram capazes de fazer cerca de 10 seqüenciamentos automatizados por dia. Em bons dias, as 230 máquinas de seqüenciamento da Celera decifravam 100 milhões de nucleotídeos de seqüências, um aumento incrível em relação às primeiras tecnologias de seqüenciamento.

Claramente, uma pequena competição tinha um efeito significativo. Em 15 de fevereiro de 2001, o International Human Genome Sequencing Consortium, apoiado por verbas públicas, publicou seu primeiro rascunho da seqüência do genoma humano na *Nature*.¹ Então, em 16 de fevereiro de 2001, o grupo privado Celera Genomics publicou seu primeiro rascunho do genoma humano na *Science*.² Os dois rascunhos eram incrivelmente consistentes (Quadro 1). O rascunho do Consortium cobria cerca de 92% do genoma e o rascunho da Celera cobria cerca de 95% do DNA eu-

cromático no genoma. A única grande surpresa foi que o genoma continha menos genes, cerca de 30.000, que o previsto – 50.000 a 100.000.

O Consortium continuou a trabalhar em uma seqüência mais completa do genoma humano e em outubro de 2004 liberou uma seqüência quase completa do genoma (cobrindo aproximadamente 99% do DNA eucromático).³ Novamente, a surpresa foi que o número estimado de genes era ainda menor, apenas 20.000 a 25.000 genes que codificam proteínas.

Agora que está disponível uma seqüência quase completa do genoma, as perguntas seguintes a serem respondidas são: “Quais são as funções de seus genes?” e “Como os elementos não gênicos influenciam a expressão desses genes?” Com a meta de responder à primeira pergunta, cientistas formaram uma International Human Proteome Organization (HUPO). Seu objetivo é determinar as estruturas e funções de todas as proteínas codificadas pelo genoma humano. Outros cientistas enfocam a segunda pergunta. Eles organizaram o ENCODE (ENCyclopedia Of DNA Elements) Consortium e tentarão identificar elementos funcionais à exceção de genes no genoma humano.

Quadro 1 Comparações Entre os Dois Primeiros Rascunhos do Genoma Humano e a Cobertura Mais Recente de 99% do DNA Eucromático

Rascunho	Pares de Nucleotídeos Seqüenciados	Número Estimado de Espaços	Número Estimado de Genes
HGP Consortium 2001	2.692.900.000	147.821	31.778
Celera Genomics 2001	2.653.979.733	105.264	26.588
HGP Consortium 2004	2.851.330.913	341	22.287

QUESTÕES PARA DISCUSSÕES

- Foram dadas patentes para proteger os investimentos de pessoas e empresas na aquisição de novos métodos e produtos. Entretanto, tem havido considerável debate sobre se os genes devem estar submetidos a proteção de patentes. Você acha que genes humanos são patenteáveis? Caso sim, a patente deve ser dada à pessoa que forneceu o DNA ou à pessoa ou empresa que seqüenciou o gene? Quais são os prós e os contras em conferir patentes de genes humanos?

¹International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. *Nature* 409: 860-921.

²Venter, J. C., et al. 2001. The sequence of the human genome. *Science* 291: 1304-1351.

- Enquanto trabalhavam em seu rascunho da seqüência do genoma humano, cientistas da Celera Genomics tiveram acesso a dados de seqüências gerados pelo International Human Genome Sequencing Consortium. Os dados do Consortium foram depositados nos bancos de dados do GenBank, EMBL e DDBJ com livre acesso a todos. Em contraste, os dados da Celera não estavam disponíveis para os cientistas do Consortium. Quais são as vantagens de tornar os dados de seqüência disponíveis grátis a qualquer um? Quais as desvantagens desta medida, especialmente quando os dados são gerados por pesquisas em empresas privadas?

³International Human Genome Sequencing Consortium. 2004. Finishing the euchromatic sequence of the human genome. *Nature* 431: 931-945.

Exercícios Básicos Ilustram a análise genética básica.

1. O que é um mapa genético?

Resposta: Um mapa genético mostra as posições de genes e outros marcadores tais como polimorfismos de comprimento de fragmentos de restrição (RFLP) em um cromossomo com base em frequências de recombinação.

2. O que é um mapa citológico?

Resposta: Um mapa citológico mostra as posições dos genes e outros marcadores genéticos em relação aos padrões de bandamento dos cromossomos.

3. O que é um mapa físico de uma molécula de DNA ou cromossomo?

Resposta: Um mapa físico de uma molécula de DNA ou cromossomo fornece as posições de genes ou outros marcadores com base nas distâncias reais em pares de bases (pb), pares de quilobases (kb) ou pares de megabases (mb), separando-os. Mapas de restrição, mapas de *contig* e sítios de seqüências marcadas (STS) são exemplos de mapas físicos.

4. Como podem ser correlacionados mapas genéticos, mapas citológicos e mapas físicos de cromossomos?

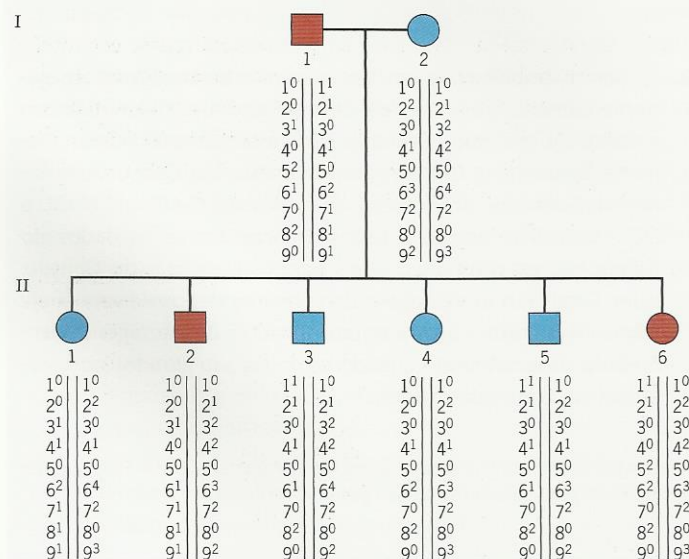
Resposta: Se um gene é clonado e posicionado em todos os três mapas, ele fornece um marcador âncora que pode ser usado para relacionar os mapas genético, citológico e físico uns com os outros. Todos os três tipos de mapas são arranjos co-lineares que mostram as localizações de seqüências de nucleotídeos no cromossomo. Eles diferem nas unidades que são usadas para situar as posições de marcadores ao longo de arranjos lineares.

5. Como a posição de mapa de um gene em um cromossomo pode ser usada para identificar e clonar o gene?

Resposta: Uma vez que um gene tenha sido posicionado no mapa genético, citológico ou físico de um cromossomo, marcadores moleculares tais como RFLP próximos ao gene podem ser usados para iniciar o andar e saltar no cromossomo, começando no marcador ligado e progredindo ao longo do cromossomo até a posição do gene de interesse. A identidade do gene deve ser estabelecida transformando um organismo mutante com uma cópia tipo selvagem do gene e mostrando que ela restaura o fenótipo tipo selvagem ou, em humanos, comparando as seqüências de nucleotídeos do gene em várias pessoas afetadas e não afetadas (veja Fig. 16.8).

Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. A doença Best é uma forma de cegueira em humanos que se desenvolve gradualmente em adultos. Ela é causada por uma mutação autossômica dominante no cromossomo 11. Nove RFLP, designados de 1 a 9, estão mapeados no cromossomo 11 em ordem numérica. Os polimorfismos em cada sítio são designados por expoentes de 0 a N, em que $N + 1$ é o número de polimorfismos presentes em um sítio na família representada pelo



heredograma desta página. O DNA foi obtido de cada membro da família, digerido com a enzima de restrição apropriada, submetido a eletroforese em gel, transferido para uma membrana de náilon por transferência de Southern, desnaturado e hibridizado com sondas radioativas que detectam todos os RFLP. Após hibridização, as membranas foram expostas a um filme de raios X e os auto-radiogramas foram usados para determinar quais RFLP estavam presentes em cada membro da família. Os resultados são mostrados no heredograma desta página. Os círculos representam mulheres; os quadrados representam homens; os símbolos vermelhos indicam pessoas com a doença Best. Qual sítio de RFLP está mais próximo da mutação que causa a doença Best? Que alelo deste RFLP está presente no cromossomo que leva a mutação da doença Best?

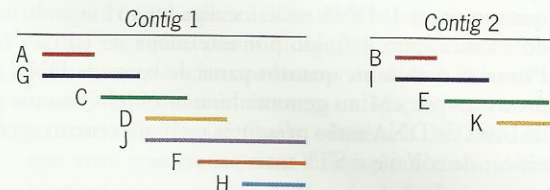
Resposta: O RFLP do sítio 4 está mais próximo da mutação da doença Best, que está presente na cópia do cromossomo 11 que leva o 4^o alelo do polimorfismo. Dos polimorfismos no cromossomo 11, apenas o 4^o alelo está presente em todos os três membros da família com a doença Best e ausente de todos os cinco membros com visão normal.

2. Onze clones genômicos, cada um contendo DNA do cromossomo 4 de *Drosophila melanogaster*, foram testados quanto a hibridização cruzada em todas as combinações pareadas. Os clones são designados de A a K, e os resultados das hibridizações são mostrados no quadro seguinte. Um sinal mais indica que ocorreu hibridização; um sinal menos indica que não foi observada hibridização.

	A	B	C	D	E	F	G	H	I	J	K
K:	-	-	-	-	+	-	-	-	-	-	+
J:	-	-	+	+	-	+	-	+	-	+	
I:	-	+	-	-	+	-	-	-	+		
H:	-	-	-	-	-	+	-	+			
G:	+	-	+	-	-	-	+				
F:	-	-	-	+	-	+					
E:	-	-	-	-	+						
D:	-	-	+	+							
C:	-	-	+								
B:	-	+									
A:	+										

Com base nos resultados de hibridização mostrados no quadro, quantos *contigs* estes clones definem? Desenhe o(s) mapa(s) de *contigs* definido(s) por estes dados.

Resposta: Os mapas dos *contigs* 2 definidos pelas 11 mutações são os seguintes:



Questões e Problemas Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 16.1. Diferencie um mapa genético, um mapa citogenético e um mapa físico. Como cada um destes tipos de mapas pode ser usado para identificar um gene por clonagem posicional?
- 16.2. Qual a diferença entre andar no cromossomo e saltar no cromossomo? Por que deve ser usado saltar no cromossomo, em lugar de andar no cromossomo, para identificar um gene de interesse?
- 16.3. O que é um *contig*? um RFLP? um VNTR? um STS? uma EST? Como cada um deles é usado na construção de mapas cromossômicos?
- 16.4. A seguinte é uma transferência de Southern de um DNA digerido de *EcoRI* de plantas de centeio de duas linhagens diferentes endocruzadas, A e B. O auto-radiograma I anexo mostra as bandas que resultam de sondar a transferência com cDNA1 marcado com ^{32}P . O auto-radiograma II mostra a mesma transferência de Southern após retirada da sonda e colocada outra sonda de cDNA2 marcada com ^{32}P .

I		II	
A	B	A	B
a1	—	b1	—
a2	—	b2	—
a3	—	b3	—
a4	—		

(a) Quais bandas você esperaria ver no auto-radiograma de uma transferência de Southern similarmente sondada preparada usando-se DNA digerido com *EcoRI* de plantas híbridas de F₁ produzidas pelo *crossing* de duas linhagens endocruzadas? (b) O que você pode concluir sobre o(s) gene(s) representado(s) pela banda a1 na transferência I nos dois endocruzamentos? (c) As plantas de F₁ foram cruzadas com plantas que possuem apenas as bandas a1, a4 e b3. O DNA foi isolado de vários indivíduos da prole e digerido com *EcoRI*. Os fragmentos de DNA resultantes foram separados por eletroforese em gel, transferidos para uma membrana de náilon

e hibridizados com sondas radioativas cDNA1 e cDNA2. O quadro seguinte resume as bandas presentes em auto-radiogramas obtidos usando-se DNA de prole individual. Interprete estes dados. Os dados fornecem evidências de RFLP? em quantos *loci*? Algum dos RFLP são ligados? Caso sim, quais são as distâncias de ligação definidas pelos dados?

Planta Nº	Bandas Presentes						
	a1	a2	a3	a4	b1	b2	b3
1	+	+	+	+			+
2	+	+	+	+			+
3	+	+	+	+			+
4	+	+	+	+			+
5	+	+	+	+	+	+	+
6	+			+	+	+	+
7	+			+	+	+	+
8	+			+	+	+	+
9	+			+	+	+	+
10	+			+			+

- 16.5. Como parte do Human Genome Mapping Project, você está tentando clonar um gene envolvido em câncer de cólon. Sua primeira etapa é localizar o gene usando marcadores de RFLP. No quadro seguinte, os *loci* de RFLP são definidos por números de STS (sítios de sequência marcada) – por exemplo, STS1 – e o gene para câncer de cólon é designado C.

<i>Loci</i>	% recombinação	<i>Loci</i>	% recombinação
C, STS1	50	STS1, STS5	10
C, STS2	15	STS2, STS3	30
C, STS3	15	STS2, STS4	14
C, STS4	1	STS2, STS5	50
C, STS5	40	STS3, STS4	16
STS1, STS2	50	STS3, STS5	25
STS1, STS3	35	STS4, STS5	41
STS1, STS4	50		

(a) Considerando a porcentagem de recombinação entre diferentes *loci* RFLP e o gene para câncer de cólon no quadro, desenhe um mapa genético mostrando a ordem e as distâncias genéticas entre marcadores RFLP adjacentes e o gene para câncer de cólon. (b) Considerando que o genoma humano contém aproximadamente $3,3 \times 10^9$ pares de bases de DNA e que o mapa genético humano contém aproximadamente 3.300 centiMorgans, aproximadamente quantos pares de DNA estão localizados ao longo do trecho do cromossomo definido por este mapa de RFLP? (Dica: Primeiro determine quantos pares de bases de DNA estão presentes por cM no genoma humano.) (c) Quantos pares de bases de DNA estão presentes na região entre o gene de câncer de cólon e o STS mais próximo?

16.6. O que são microssatélites? Por que são chamados de microssatélites?

16.7. Você clonou um gene humano previamente desconhecido. Que procedimento lhe permitiria posicionar este gene no mapa citológico do genoma humano sem fazer nenhuma análise de heredograma? Descreva como você faria este procedimento.

16.8. Você identificou uma EST humana previamente não descrita. O que deve ser feito antes que esta nova EST possa ser chamada de STS?

16.9. VNTR e microssatélites são classes específicas de polimorfismos. Qual a diferença entre um VNTR e um microssatélite? Por que seria impróprio chamar a maioria dos VNTR de microssatélites?

16.10. Um RFLP e um alelo mutante que causa albinismo em humanos não pode ser mostrado como separado por recombinação com base na análise de heredogramas ou mapeamento híbrido por radiação. Estas observações significam que o RFLP ocorre dentro do gene que contém a mutação causadora do albinismo ou se superpõe a ele? Caso sim, por quê? Caso não, por que não?

16.11. Um RFLP e uma mutação que causa surdez em humanos estão ambos mapeados no mesmo local no mesmo cromossomo. Como você pode determinar se há ou não superposição do RFLP com o gene contendo a mutação de surdez?

16.12. Por que a resolução do mapeamento de híbridos por radiação de genes humanos é maior do que a resolução do mapeamento padrão de híbridos de células somáticas?

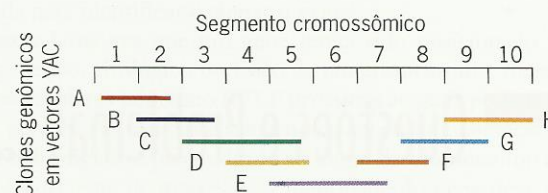
16.13. Quais são as metas do Human Genome Project? Que impacto essas metas terão na prática da medicina no futuro? Quais são os possíveis usos inapropriados de dados do genoma humano?

16.14. Para ser útil como um marcador genético para clonagem posicional de um alelo mutante que causa uma anomalia herdada em humanos, um RFLP deve estar presente no mesmo homólogo que a mutação. Por quê?

16.15. Que tipo de marcador molecular, RFLP ou EST, é mais provavelmente um marcador de gene mutante causador de doença em humanos? Por quê?

16.16. O bacteriófago ΦX174 contém 11 genes em um genoma de 5.386 pb; *E. coli* tem 4.288 genes previstos em um genoma de cerca de 4.639 kb; *S. cerevisiae* tem cerca de 6.000 genes em um genoma com tamanho de 12.1 mb; *C. elegans* tem cerca de 19.000 genes presentes em um genoma de cerca de 100 mb; e *H. sapiens* tem um número estimado de 22.000 genes em seu genoma de 3.000 mb. Que genoma tem a maior densidade gênica? a menor densidade gênica? Parece haver alguma correlação entre densidade gênica e complexidade desenvolvimental? Caso sim, descreva a correlação.

16.17. É mostrado um mapa de contig de um segmento do cromossomo 3 de *Arabidopsis*.



(a) Se uma EST hibridiza com os clones genômicos C, D e E mas não com os outros clones, em que segmento do cromossomo 3 está localizada a EST? (b) Se um clone do gene *ARA* só hibridiza com os clones genômicos C e D, em que segmento cromossômico o gene está localizado? (c) Se um fragmento de restrição hibridiza só com um dos clones genômicos mostrados previamente, em que segmento(s) cromossômico(s) o fragmento estaria localizado?

16.18. Oito híbridos de radiação humano-hamster chinês foram testados quanto à presença de seis EST designadas de A a F. Os resultados são mostrados no seguinte quadro, onde um mais indica que o marcador está presente e um menos indica que está ausente.

	Híbrido de radiação							
	1	2	3	4	5	6	7	8
A	-	+	-	-	+	+	-	+
B	+	-	+	-	-	+	-	+
C	-	+	+	+	-	-	-	+
D	+	-	+	+	-	-	+	-
E	+	-	+	+	-	-	+	-
F	-	+	-	+	+	+	+	-

Com base nestes dados, algumas das EST parecem estar proximamente ligadas? Quais? O que seria necessário para você ter mais certeza de sua resposta?

16.19. Que vantagens *chips* gênicos têm em relação a técnicas de hibridização tradicionais de *dot blot*?

16.20. Que vantagem importante a proteína fluorescente verde da água-viva tem em relação a outros métodos para estudar a síntese e a localização de proteínas?

16.21. Você recebe bibliotecas de cDNA específicas de cromossomos para todos os 24 cromossomos humanos. Como essas bibliotecas podem ser usadas para estudar a evolução cromossômica em primatas?

- 16.22. Das espécies de cereais, apenas o milho contém duas cópias de cada bloco de genes ligados. O que esta duplicação de conjuntos de genes humanos indica sobre a origem desta espécie agronomicamente importante?
- 16.23. Cinco clones de DNA genômico humano presentes em vetores YAC foram testados por hibridização quanto à presença de seis sítios de sequência marcada designados STS1 até STS6. Os resultados são dados no quadro seguinte; um sinal mais indica a presença de STS e um menos indica a ausência de STS.

		STS					
		1	2	3	4	5	6
Clone de YAC	A	+	-	+	+	-	-
	B	+	-	-	-	+	-
	C	-	-	+	+	-	+
	D	-	+	-	-	+	-
	E	-	-	+	-	-	+

- (a) Qual a ordem de sítios STS no cromossomo?
 (b) Desenhe o mapa de *contig* definido por estes dados.

- 16.24. No começo deste capítulo, discutimos o *pool* gênico único na Islândia e a pesquisa da deCODE Genetics enfocada na identificação de genes associados a doenças humanas. Um dos genes estudados pelos cientistas da deCODE codifica a proteína ativadora de 5-lipoxigenase (FLAP), que está envolvida na síntese de moléculas sinalizadoras nos glóbulos brancos. Use o *web site* PubMed Entrez (URL: <http://www.ncbi.nlm.nih.gov/entrez>) para procurar informação sobre o trabalho da deCODE relativo ao gene codificante de FLAP. *Dica*: encurte sua pesquisa usando uma entrada (*query*) de duas partes: "5-lipoxigenase activating protein" e "deCODE." (Os dois componentes da entrada devem ser separados por uma vírgula seguida de um espaço.) Que doenças humanas são afetadas por alelos diferentes do gene humano codificante desta proteína?

- 16.25. Suponha que você sequenciou um pequeno fragmento de DNA que você clonou. A sequência de nucleotídeos deste segmento de DNA é mostrada em seguida.

aagtagtcgaaaccgaattccgtagaacaactcgacgctccggttcgtg
 ttgcaacaaatagcattccatcgcgagcttagaatcaccgagtgccaga
 gtcacgttcgtaagcagggcagtttacaggcagcagaaaaatcgattgaaca
 gaaatggctggcggttaagcagggcaaggattcgggcaaggccaaggcgaagg
 cggtagtcggttcgcgcgcgagg

Em uma tentativa de aprender algo sobre a identidade ou possível função desta sequência de DNA, você decide fazer uma pesquisa BLAST (megablast) no *web site* PubMed Entrez (vide problema anterior). Digite ou arraste esta sequência para o box de sequência indagada e faça as mesmas alterações no Format que foram sugeridas no Enfoque Técnico: GenBank para limitar a quantidade de informações obtidas. Faça a pesquisa e examine as sequências mais proximamente relacionadas à sua indagada. Elas são sequências codificantes? Que proteínas codificam? A primeira sequência listada é NM_079795.2. Vá para o

ENTREZ "Search across databases" (clique no box na parte inferior da *home page* Entrez), digite ou cole NM_079795.2 como indagação e clique em "Go". Você terá os resultados em três bancos de dados: nucleotídeo, HomoloGene e UniGene. Examine a informação em todos os três bancos de dados e veja o que você pode aprender sobre a função de seu DNA ou da proteína que ele codifica. O que você aprendeu sobre sua sequência de DNA destas pesquisas? Repita a procura de megablast com apenas metade de sua sequência como indagação. Você ainda identifica as mesmas sequências nos bancos de dados? Se você usar um quarto de sua sequência como indagação, você ainda obtém a mesma sequência? Qual a sequência de DNA mais curta que você pode usar como indagação e ainda identificar as mesmas sequências nos bancos de dados?

- 16.26. O *web site* Entrez do PubMed (veja Problema 16.24) também pode ser usado para procurar sequências de proteínas. Em vez de fazer uma pesquisa "megablast" com um ácido nucleico, faz-se um "blastp" com um polipeptídeo (sequência de aminoácidos). Suponha que você tem a seguinte sequência parcial de um polipeptídeo:

GYDVEKNNSRIKLGKSLVSKGILVQTKGTGASG
 SFKLNKKAASGEAKPQAKKAGAAKA

Vá para o *web site* Entrez e clique protein no topo da barra de ferramentas, então clique BLAST à esquerda. Clique então em protein-protein BLAST (blastp) e entre com sua sequência indagada no box do topo. Role para "Format" e mude de "HTML" para "Plain text". Mude também o número de "Descriptions" e "Alignments" ambos para 10. Clique então em "BLAST!" e "Format" na página subsequente. Seus resultados devem estar disponíveis em 10 a 15 segundos. A primeira sequência listada deve ser idêntica a sua sequência indagada; a segunda sequência deve diferir de sua indagação por um único aminoácido. Qual a identidade de sua sequência indagada?

- 16.27. A sequência de um gene em *Drosophila melanogaster* que codifica um polipeptídeo da histona H2A é:

aagtagtcgaaaccgaattccgtagaacaactcgacgctccggttcg
 tgttgcaacaaaatagcattccatcgcgagcttagaatcaccgagtg
 cccagagtcacgttcgtaagcagggcagtttacaggcagcagaaaaa
 tcgattgaacagaaatggctggcggttaagcagggcaaggattcgggcaa
 ggccaaggcgaaggcggtatcgcttcgcgcgcggttccttcagtt
 ccccggtgggtcgatccatcgctatcgaagcggcactacgtcacatgg
 acggtcgagcgaactcgacgctgtactcggctgcatattggaatacctg
 accgccgaggtcctggagttggcaggaacgcacgaaggacttgaaag
 tgaacgtatcactctcgccactacagctcgccattcgaggagacgag
 gagctggacagcctgatcaaggcaaccatcgctggtggcggtgctatcc
 gcacatacacaagtcgctgatcggaacaaaggaggaaacggtgca
 ggatccgcagcgggaagggaacgtattctgctgcaggcctactaagcca
 gtcggcaatcggacgcttcgaaacatgcaacactaatgtttaattca
 gatttcagcagagacaagctaaaacacgcagcaggttgtaattctctg
 tgcgccagcatatattcttatatacaacgtaatacataattatgtaattc
 tagcatctcccaacacacacatacaaaaaaatacaaacacacaaaa
 cgtatttaccgcagcagcattcctggcgaggttgtagatgaacaaaaaact
 taatttagagcaaaagtaattacgaataaattataaaaaaactataataa
 aacgcc.

Usemos o programa de tradução disponível na Internet em <http://www.expasy.org/tools/dna.html> para traduzir este gene em todas as seis matrizes de leitura possíveis e ver qual matriz de leitura específica a histona H2A. Digite ou cole a sequência de DNA no box “ExPASy Trans-

late Tool” e clique “TRANSLATE SEQUENCE.” Os resultados mostrarão os produtos da tradução em todas as seis matrizes de leitura com **Met** e **Stop** em negrito para destacar as potenciais matrizes de leitura. Qual delas especifica a histona H2A?

Aplicações da Genética Molecular

Tópicos

- Uso da Tecnologia de DNA Recombinante para Identificar Genes Humanos
- Diagnóstico Molecular de Doenças Humanas
- Terapia Gênica Humana
- Fingerprints de DNA
- Produção de Proteínas Eucarióticas em Bactérias
- Plantas e Animais Transgênicos
- Genética Reversa: Dissecando Processos Biológicos por Inibição da Expressão Gênica



Um pré-embrião humano de oito células.

Detecção da Mutação Tay-Sachs em Pré-embriões de Oito Células

Brittany foi concebida *in vitro* a partir de um ovócito e um espermatozóide obtidos de sua mãe e seu pai. Três dias após a concepção, uma das oito células do pré-embrião do qual Brittany se desenvolveu foi usada para fazer testes de DNA em busca de um gene mutante que causa a doença Tay-Sachs. Os resultados dos testes foram negativos, e o pré-embrião foi implantado no útero da mãe de Brittany. Nove meses depois, nasceu Brittany sem a preocupação de que algum dia ela poderia morrer da doença Tay-Sachs, como sua irmã mais velha. A remoção de uma célula de um pré-embrião de oito células não afeta seu desenvolvimento subsequente.

A **doença Tay-Sachs** (Veja Enfoque Técnico: Doença de Tay-Sachs, uma Tragédia na Infância, no Cap. 13) é um distúrbio autossômico recessivo letal. Crianças com Tay-Sachs são normais ao nascimento, mas sofrem rápida degeneração neurológica que evolui para cegueira, paralisia, retardo mental e morte aos três a quatro anos de idade. A degeneração resulta da ausência de uma enzima cha-

mada hexosaminidase A, que catalisa a primeira etapa na degradação de um lipídeo complexo chamado gangliosídeo G_{M2} . Na ausência de hexosaminidase A, o gangliosídeo G_{M2} acumula-se nos neurônios e causa degeneração progressiva do sistema nervoso central.

Ambos os pais de Brittany são heterozigotos para o gene mutante que causa doença Tay-Sachs. Eles sabiam que, se tivessem outro filho, haveria uma chance de $\frac{1}{4}$ de ele sofrer a doença Tay-Sachs, como sua primeira filha. Embora eles quisessem um filho, os genitores de Brittany não queriam correr o risco de ter outro filho com esta doença. A doença Tay-Sachs pode ser detectada pré-natalmente. Entretanto, por motivos religiosos, o casal havia excluído o aborto de um feto acometido. A fertilização *in vitro* e um teste de DNA para o gene mutante de hexosaminidase A que ambos os genitores fizeram permitiu a confirmação de que Brittany não era homozigota para o gene mutante antes da implantação ainda no estágio de pré-embrião. Para seus genitores, que sofreram com a degeneração e a morte do primeiro filho, Brittany era um tesouro sem preço.

Os testes de DNA para Tay-Sachs e outros distúrbios humanos foram possibilitados com a descoberta de defeitos genéticos que causam tais doenças. Os defeitos foram identificados pesquisando-se o genoma humano com o uso de técnicas sofisticadas de mapeamento cromossômico e DNA recombinante. Neste capítulo, consideraremos algumas das muitas aplicações do DNA recombinante, clonagem de genes e tecnologias genômicas descritas nos dois capítulos anteriores. Essas tecnologias são, sem dúvida, as ferramentas mais poderosas já desenvolvidas no campo da biologia.

Do mesmo modo que os geneticistas hoje conhecem toda a via de morfogênese do bacteriófago T4 (Cap. 13), no futuro eles saberão a via completa de morfogênese de uma célula de levedura, uma mosca-das-frutas, uma planta *Arabidopsis* ou mesmo um ser humano. Além disso, em algum ponto, os biólogos compreenderão a base molecular de aprendizagem e memória, e saberemos que eventos moleculares estão subjacentes ao processo de envelhecimento. Mais importante, eles compreenderão os mecanismos complexos que regulam a divisão celular em humanos e deverão ser capazes de usar esse conhecimento para evitar ou curar pelo menos alguns tipos de câncer humano e infecções virais ameaçadoras à vida.

USO DA TECNOLOGIA DE DNA RECOMBINANTE PARA IDENTIFICAR GENES HUMANOS

Os genes mutantes que causam a doença Huntington e a fibrose cística foram identificados por clonagem posicional.

As técnicas de DNA recombinante revolucionaram a pesquisa à procura de genes defeituosos que causam

doenças humanas. Vários “genes de doença” já foram identificados por clonagem posicional (Cap. 16). Além disso, as mutações responsáveis pelas doenças foram determinadas comparando-se as seqüências de nucleotídeos de alelos tipo selvagem e mutantes dos genes. As seqüências codificantes dos alelos tipo selvagem foram traduzidas por computadores para prever as seqüências de aminoácidos dos produtos gênicos. Oligopeptídeos foram sintetizados com base nas seqüências previstas de aminoácidos e usados para produzir anticorpos, que, por sua vez, foram usados para localizar os produtos gênicos e investigar suas funções *in vivo*. Os resultados desses estudos permitirão futuros tratamentos de algumas destas doenças por terapia gênica.

Doença Huntington

A doença Huntington (HD) é um distúrbio insidioso causado por uma mutação autossômica dominante, que ocorre em cerca de uma em cada 10.000 pessoas descendentes de europeus. As pessoas com HD sofrem degeneração progressiva do sistema nervoso central, geralmente começando entre os 30 a 50 anos e terminando em morte 10 a 15 anos depois. Hoje em dia, a HD é intratável. Entretanto, a identificação do gene e do defeito mutacional responsável pela HD deu esperanças de um tratamento eficaz no futuro. Devido à idade tardia de manifestação da doença, a maioria dos pacientes com HD já tiveram filhos antes de aparecerem os sintomas da doença. Como o distúrbio é causado por uma mutação dominante, cada filho de um paciente heterozigoto para HD tem uma chance de 50% de ser afetado pela doença. Essas crianças observam a degeneração e morte de seu genitor com HD, sabendo que têm uma chance de 50% de ter o mesmo destino.

O gene HD foi um dos primeiros genes humanos que se demonstrou estarem fortemente ligados a um RFLP. Em 1983, James Gusella, Nancy Wexler e colaboradores demonstraram que o gene da HD segregava junto com um RFLP mapeado perto da ponta do braço curto do cromos-

somo 4. Eles basearam seus achados amplamente em dados de estudos de duas grandes famílias, uma na Venezuela e uma nos EUA. Pesquisas subsequentes mostraram que a ligação era cerca de 96% completa; 4% da prole de heterozigotos HD eram recombinantes para o RFLP e o alelo de HD. Considerando esta localização inicial do gene da HD em um segmento relativamente pequeno do cromossomo 4, alguns geneticistas previram que o gene da HD logo seria clonado e caracterizado. Entretanto, a tarefa foi mais difícil do que o esperado e levou 10 anos para ser realizada.

Usando procedimentos de clonagem posicional, Gusella, Wexler e colaboradores identificaram um gene, primeiro chamado *IT15* (de Transcrito de Interesse número 15) e subsequentemente denominado *huntingtina*, que ocupa cerca de 210 kb perto da ponta do braço curto do cromossomo 4 (Fig. 17.1). Este gene contém uma repetição de trinucleotídeo, (CAG)_n, que está presente em 11 a 34 cópias em cada cromossomo 4 de pessoas saudáveis. Nas pessoas com HD, o cromossomo que tem a mutação HD contém 42 a mais de 100 cópias da repetição CAG neste gene. Além disso, a idade de início de HD é inversamente correlacionada com o número de cópias da repetição do trinucleotídeo. Início juvenil raro da doença ocorre em crianças com um número de repetições de cópia incomumente alto. As regiões de repetição do trinucleotídeo de cromossomos HD são instáveis, com os números de repetições em geral aumentando e às vezes diminuindo entre gerações. Gusella, Wexler e colaboradores detectaram regiões de repetição CAG ampliadas em cromossomos de 72 famílias diferentes com HD, deixando poucas dúvidas de que haviam identificado o gene correto.

O gene *huntingtina* é expresso em muitos tipos diferentes de células, produzindo um grande mRNA com 10 a 11 kb. A região codificante do mRNA de *huntingtina* prevê uma proteína com 3.144 aminoácidos de tamanho. Infelizmente, a sequência prevista de aminoácidos da proteína huntingtina fornece poucas informações sobre sua função. A dominância da mutação HD indica que a proteína mutante causa a doença.

A região de repetição CAG ampliada no gene mutante *huntingtina* codifica uma região anormalmente longa de poliglutamina perto do terminal amino da proteína. A região alongada de poliglutamina favorece interações proteína-proteína que resultam no acúmulo de agregados da proteína huntingtina nas células do cérebro. Esses agregados de proteína são tidos como causadores dos sintomas clínicos de HD, e enfoques atuais de tratamento envolvem tentativas de desarticular ou eliminar esses agregados de proteínas.

HD foi a quarta doença humana a ser associada a uma repetição instável de trinucleotídeos. Em 1991, a síndrome do X frágil – a forma mais comum de retardo mental em humanos – foi o primeiro distúrbio humano a ser associado a uma repetição expandida de trinucleotídeos. Nós discutimos a descoberta da natureza da mutação reponsável pela síndrome do X frágil em Um Marco na Genética: Repetições de Trinucleotídeos e Doença Humana. Pouco depois, foi demonstrado que a distrofia miotônica e a atrofia muscular espinobulbar (ambas doenças associadas a perda de controle muscular) resultam de repetições expandidas de trinucleotídeos. Em 2004, foi demonstrado que mais de 40 distúrbios humanos diferentes, muitos associados a anomalias neurodegenerativas, resultam de repetições ampliadas de trinucleo-

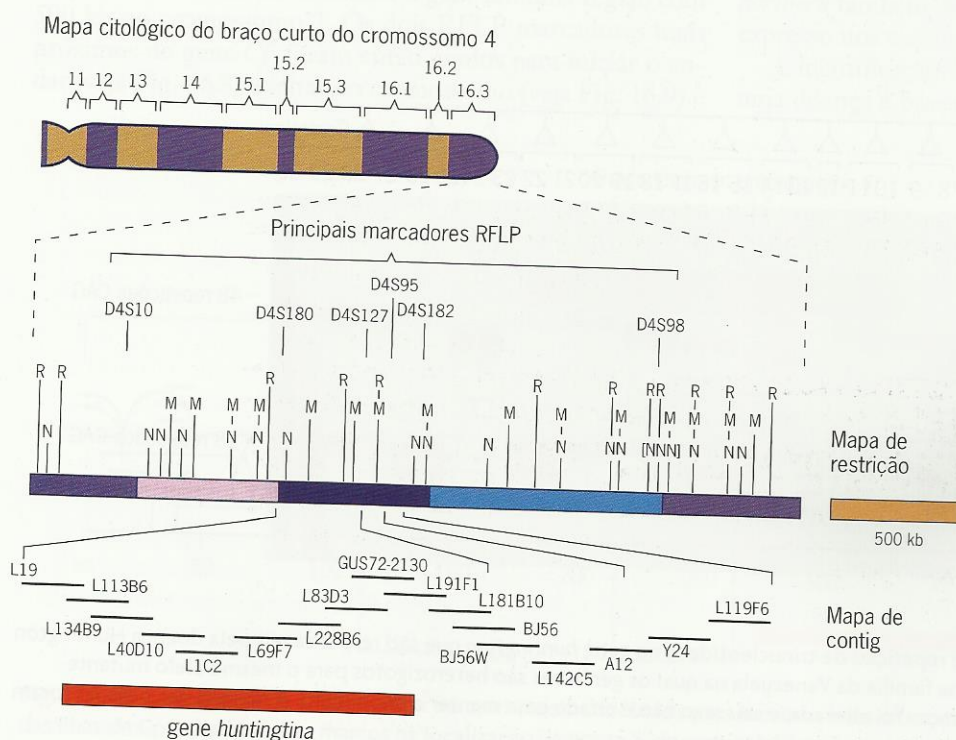


Fig. 17.1 ■ Identificação do gene responsável pela doença Huntington por clonagem posicional. O mapa citológico do braço curto do cromossomo 4 é mostrado no topo. Os marcadores RFLP, mapa de restrição e mapa contig usados para localizar o gene *huntingtina* são mostrados abaixo do mapa citológico. M, N e R representam sítios de restrição MluI, NotI e NruI, respectivamente.

tídeos. Elas incluem vários tipos de ataxia espinocerebelar, atrofia palidoluisiana dentatorrubral (síndrome Haw River), ataxia de Friedreich e síndrome do X frágil. A alta frequência de distúrbios humanos causados pela expansão de repetições de trinucleotídeos indica que este pode ser um evento mutacional comum em nossa espécie.

Embora a identificação do defeito genético, a repetição expandida de trinucleotídeos no gene *huntingtina*, não tenha resultado em um tratamento do distúrbio, ela forneceu um teste simples e preciso do DNA para a mutação HD (Fig. 17.2). Uma vez conhecidas as seqüências do gene *huntingtina* em ambos os lados da região de repetição de trinucleotídeos, foi possível sintetizar *primers* oligonucleotídicos e usá-los para amplificar a região por PCR, e o número de repetições CAG foi determinado por eletroforese em gel de poliacrilamida. Assim, pessoas em risco de levar o gene mutante *huntingtina* podem ser facilmente testadas quanto a sua presença. Como o procedimento de PCR requer pouco DNA, o teste para HD também pode ser feito pré-natalmente em células fetais obtidas por amniocentese ou punção de vilosidades coriônicas (veja Enfoque Técnico: Detecção de Aneuploidia em Fetos Humanos, no Cap. 6).

Considerando a disponibilidade do teste de DNA para a mutação HD, pessoas que estão em risco de transmitir o gene defeituoso para seus filhos podem determinar se eles o possuem antes de começar uma família. Cada pessoa com um genitor heterozigoto tem uma chance de 50% de não ter o gene defeituoso. Se o teste for negativo, ela ou ele pode começar uma família sem a preocupação de transmitir a mutação. Se o teste for positivo, o feto pode ser testado pré-natalmente, ou o casal pode considerar a fertilização *in vitro*, como fizeram os genitores que discutimos no começo deste capítulo. Se o teste do embrião de oito células for negativo para a mutação HD, ele pode ser implantado no útero da mãe sabendo-se que ele tem duas cópias normais do gene *huntingtina*. Se usado conscientemente, o teste de DNA para a mutação HD deve diminuir o sofrimento humano decorrente desta terrível doença.

Fibrose Cística

Fibrose cística (CF) é uma das doenças herdadas mais comuns em humanos, afetando 1 em 2.000 neonatos de descen-

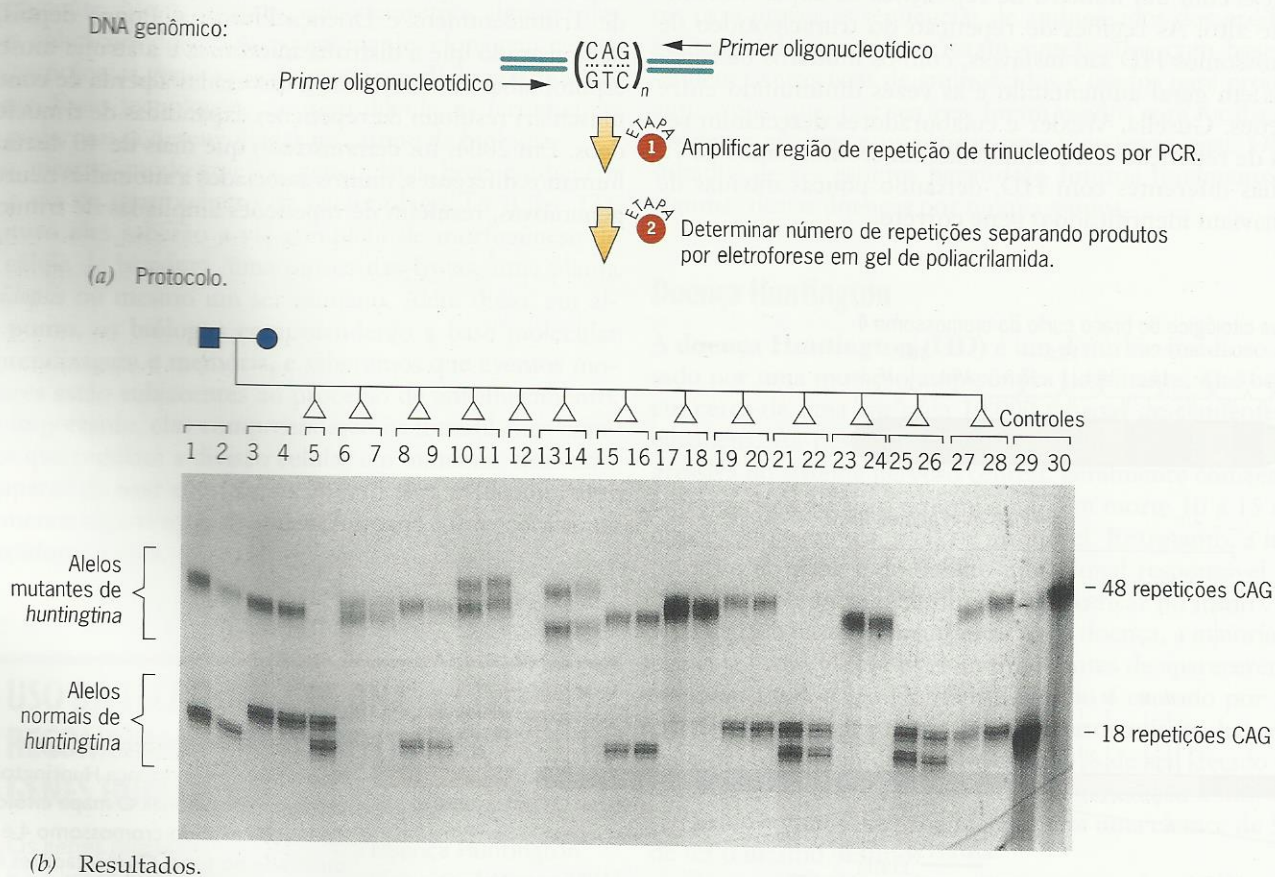


Fig. 17.2 ■ Teste das regiões expandidas de repetição de trinucleotídeos no gene *huntingtina* que são responsáveis pela doença Huntington por PCR. Os resultados mostrados são de uma família da Venezuela na qual os genitores são heterozigotos para o mesmo alelo mutante *huntingtina*. A ordem de nascimento das crianças foi alterada, e seu sexo não é citado para manter o anonimato. A maioria das pessoas foram testadas duas vezes para minimizar erros.

dência européia. A CF é herdada como uma mutação autosômica recessiva, e a frequência de heterozigotos é estimada em cerca de 1 em 25 nas populações caucasianas. Só nos EUA, mais de 30.000 pessoas sofrem desta doença devastadora. Um sintoma facilmente diagnosticado de CF é um suor excessivamente salino, um efeito benigno do gene mutante. Os outros sintomas não são nada benignos. Pulmões, pâncreas e fígado tornam-se obstruídos com um muco espesso, que resulta em infecções crônicas e por fim em mau funcionamento destes órgãos vitais. Além disso, muco geralmente se acumula no trato digestivo, fazendo com que as pessoas fiquem mal alimentadas, não importando o quanto comam. As infecções pulmonares são recorrentes, e os pacientes em geral morrem de pneumonia ou outras infecções do sistema respiratório. Em 1940, a expectativa média de vida de um neonato com CF era de menos de dois anos. Com a melhoria dos métodos de tratamento, ela foi gradualmente aumentada. Hoje em dia, a expectativa de vida para uma pessoa com CF é de cerca de 32 anos, mas a qualidade de vida é inadequada.

A identificação do gene *CF* é um dos principais sucessos da clonagem posicional (Cap. 16). A análise bioquímica de células de pacientes com CF falhou em identificar qualquer defeito metabólico específico ou produto gênico mutante. Então, em 1989, Francis Collins e Lap-Chee Tsui e seus colaboradores identificaram o gene *CF* e caracterizaram algumas das mutações que causam esta doença trágica. A clonagem e o sequenciamento do gene *CF* rapidamente resultaram na identificação de seu produto, o que então sugeriu enfoques para o tratamento clínico da doença e esperança para uma terapia gênica bem-sucedida no futuro.

O gene *CF* foi primeiro mapeado no braço longo do cromossomo 7 por sua co-segregação com RFLP. O posterior mapeamento de RFLP localizou o gene em uma região com 500 kb no cromossomo 7. Os dois RFLP marcadores mais próximos do gene *CF* foram então usados para iniciar o andar (veja Fig. 16.8) e saltar no cromossomo (veja Fig. 16.9) e

para começar a construção de um mapa físico detalhado da região (Fig. 17.3). Três tipos de informação foram usados para estreitar a pesquisa do gene *CF*.

1. Os genes humanos em geral são precedidos de grupos de citosinas e guaninas chamados de **ilhas de CpG** (Cap. 21). Três destes grupos estão presentes antecedendo o gene *CF* (Fig. 17.3).
2. Sequências codificantes importantes em geral são conservadas em espécies correlatas. Quando foram usadas sequências éxon do gene *CF* como sondas de transferências de Southern contendo fragmentos de restrição de DNA genômicos humanos, de camundongo, de *hamster* e de bovinos (em geral, chamadas de **zoo blots**), foi observado que os éxons são altamente conservados.
3. Como já mencionado, sabe-se que a CF está associada a um muco anormal nos pulmões, no pâncreas e em glândulas sudoríparas. Foi preparada uma biblioteca de cDNA a partir de mRNA isolado de culturas de células de glândulas sudoríparas e triadas por hibridização de colônia usando sondas de éxons do gene *CF* (na época o gene *CF* candidato).

O uso da biblioteca de cDNA de glândulas sudoríparas demonstrou-se crítico na identificação do gene *CF*, pois experimentos com transferência *northern* subsequentemente mostraram que este gene é expresso apenas em células epiteliais dos pulmões, do pâncreas, das glândulas salivares, das glândulas sudoríparas, do intestino e das vias reprodutivas. Assim, clones de cDNA do gene *CF* não seriam identificados usando-se bibliotecas de cDNA preparadas a partir de outros tecidos e órgãos. Os resultados das transferências *northern* também demonstraram que o hipotético gene *CF* é expresso nos tecidos apropriados.

A identificação de um gene candidato como o gene de uma doença é baseada em comparações dos alelos normal e

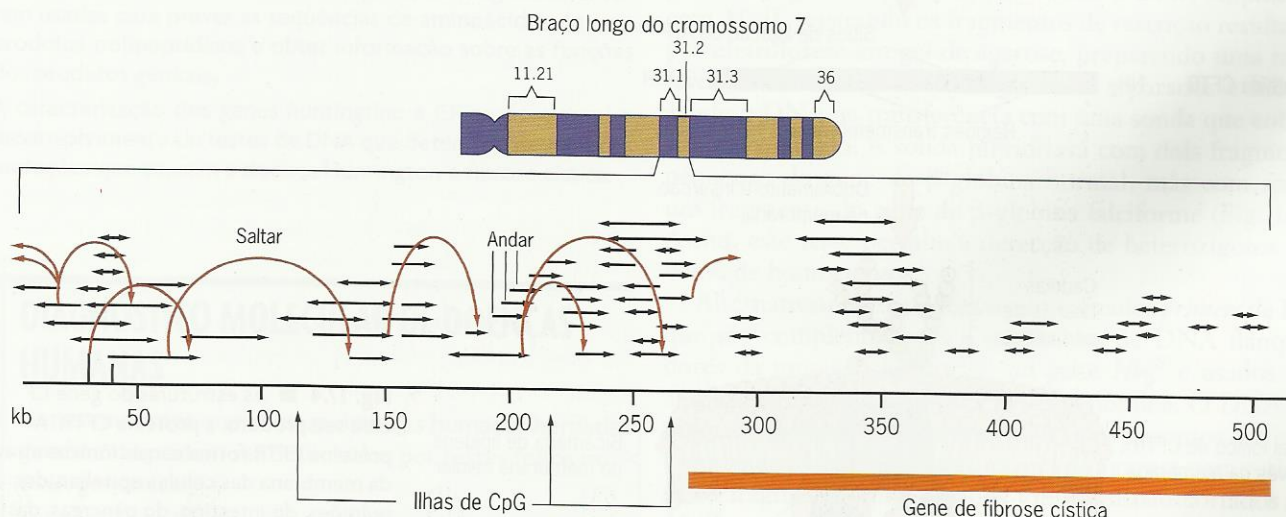


Fig. 17.3 ■ A sequência de andar e saltar no cromossomo foi usada para localizar e caracterizar o gene da fibrose cística. As posições das ilhas de CpG usadas como marcos na localização da ponta 5' do gene também são mostradas.

mutante de várias famílias diferentes. CF é incomum, pois 70% dos alelos mutantes contêm a mesma deleção de três bases, $\Delta F508$, que elimina a fenilalanina na posição 508 no produto do gene *CF*. Ao contrário do gene *huntingtina*, a sequência de nucleotídeos do gene *CF* mostrou ser informativa. O gene é grande, com 250 kb e contendo 24 éxons (Fig. 17.4). O mRNA de *CF* tem cerca de 6,5 kb de tamanho e codifica uma proteína com 1.480 aminoácidos. Uma pesquisa computadorizada dos bancos de dados de proteínas mostrou rapidamente que o produto do gene *CF* é similar a várias proteínas de canais iônicos, que formam poros entre as células através dos quais passam os íons. O produto do gene *CF*, chamado **regulador de condutância transmembranar da fibrose cística**, ou **proteína CFTR**, forma canais iônicos (Fig. 17.4) pelas membranas das células que revestem o trato respiratório, o pâncreas, as glândulas sudoríparas, o intestino e outros órgãos e regula o fluxo de sais e água para dentro e para fora destas células. Como a proteína CFTR mutante não funciona apropriadamente em pacientes com CF, o sal acumula-se nas células epiteliais e o muco acumula-se na superfície destas células.

A presença de muco no revestimento do trato respiratório resulta em infecções crônicas progressivas por *Pseudomonas aeruginosa*, *Staphylococcus aureus* e bactérias correlatas. Essas infecções, por sua vez, freqüentemente resultam em insuficiência respiratória e morte. Entretanto, as mutações no gene *CF* são pleiotrópicas. Elas causam vários efeitos distintos pleiotrópicos. Funcionamento inadequado do pâncreas, do fígado, dos ossos e do trato intestinal é comum em pessoas com CF. Embora CFTR forme canais de cloro (Fig. 17.4), ele também regula a atividade de vários outros sistemas de transporte tais como canais de potássio e de sódio. Alguns trabalhos sugerem que CFTR pode ter um papel na regulação do metabolismo e do transporte de lipídeos. O CFTR interage com várias outras proteínas e sofre fosforilação/desfosforilação por cinases e fosfatases. Assim, CFTR pode ser considerado multifuncional. De fato, alguns dos sintomas de CF podem resultar da perda de funções do CFTR à exceção dos canais de cloro.

Embora 70% dos casos de CF sejam devidos a deleção de trinucleotídeos $\Delta F508$, mais de 170 mutações *CF* diferentes já foram identificadas (Fig. 17.5). Cerca de 20 dessas mutações

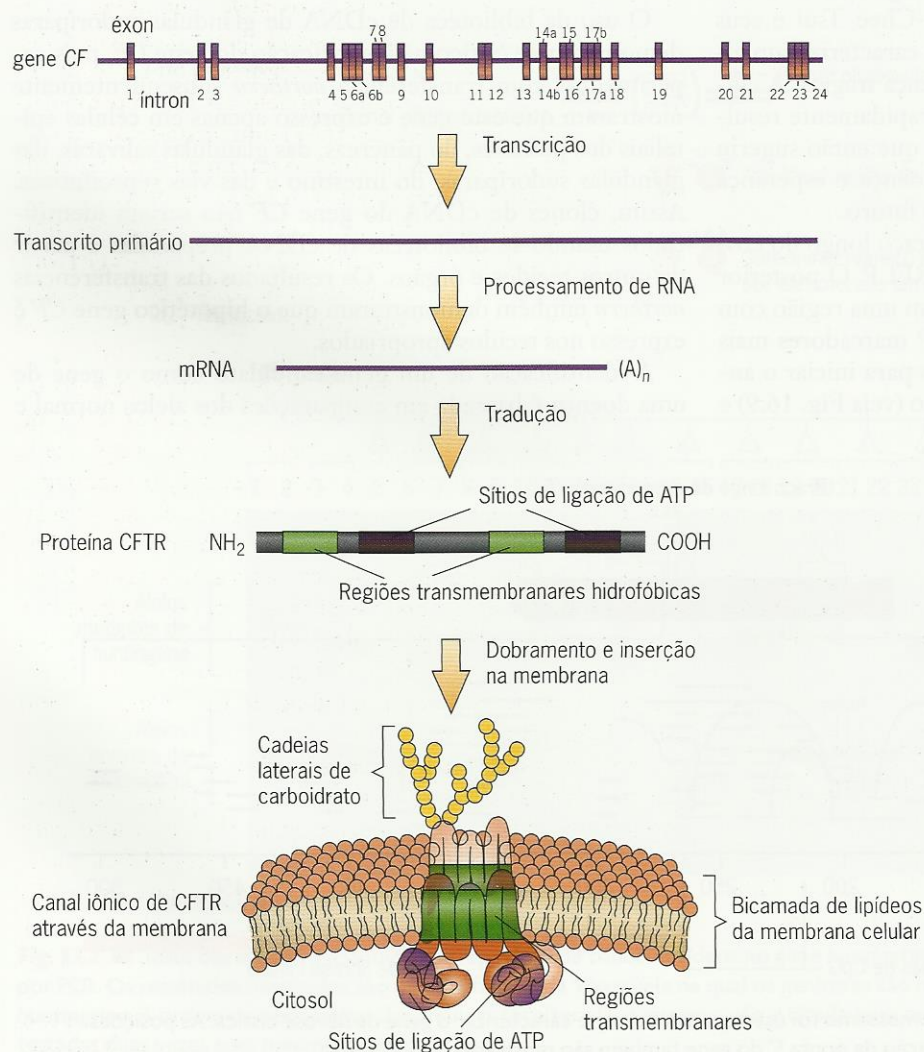


Fig. 17.4 ■ As estruturas do gene *CF* e de seu produto, a proteína CFTR. A proteína CFTR forma canais iônicos através da membrana das células epiteliais dos pulmões, do intestino, do pâncreas, das glândulas sudoríparas e de alguns outros órgãos.

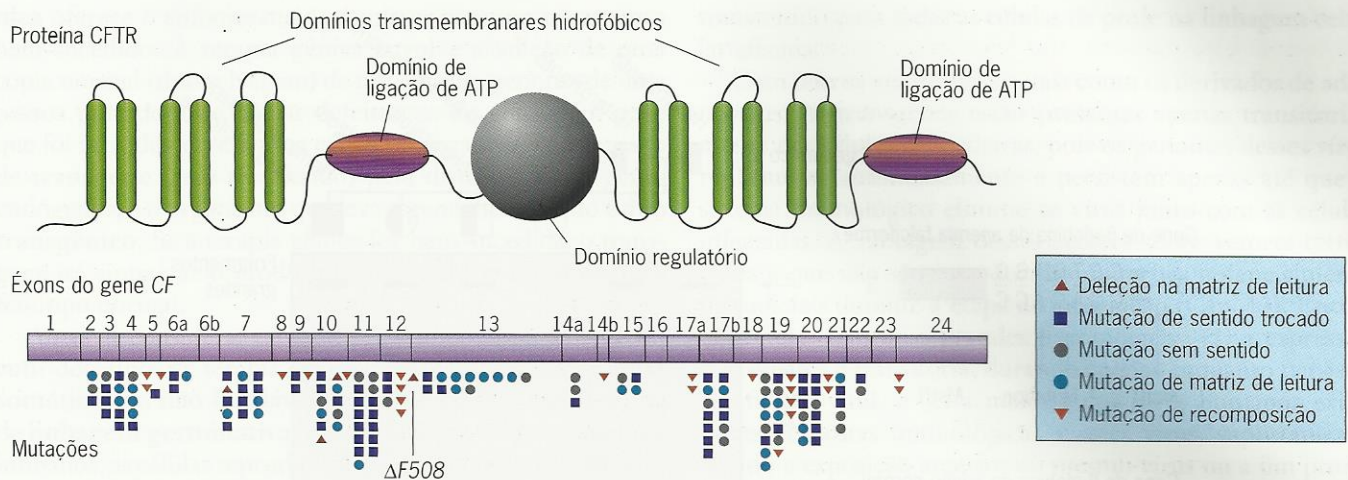


Fig. 17.5 ■ Mutações no gene CF que causam fibrose cística. A distribuição e classificação das mutações que causam fibrose cística são mostradas abaixo dos éxons do gene CF. É mostrado um diagrama esquemático da proteína CFTR acima do mapa de éxons para ilustrar os domínios da proteína que estão alterados pelas mutações. Cerca de 70% de todos os casos de CF resultam da mutação $\Delta F508$, que deleta a fenilalanina presente na posição 508 da proteína normal CFTR.

são muito comuns; outras são raras, e muitas só foram identificadas em uma só pessoa. Várias dessas mutações podem ser detectadas por triagem de DNA, tais como o teste para a deleção $\Delta F508$ ilustrada em Enfoque Técnico: Detecção de um Gene Mutante Causador de Fibrose Cística, no Cap. 15. Estes testes podem ser feitos em células fetais obtidas por amniocentese ou punção de vilosidades coriônicas. Eles também foram feitos de maneira bem-sucedida em embriões de oito células antes da implantação produzidos por fertilização *in vitro*. A diversidade de mutações que causam CF (veja Fig. 17.5) dificulta criar testes de DNA para todos os alelos mutantes CF.

PONTOS IMPORTANTES

- Os genes mutantes responsáveis pela doença Huntington e pela fibrose cística foram identificados por clonagem posicional.
- As seqüências de nucleotídeos dos genes *huntingtina* e CF foram usadas para prever as seqüências de aminoácidos de seus produtos polipeptídicos e obter informação sobre as funções dos produtos gênicos.
- A caracterização dos genes *huntingtina* e CF proporcionou o desenvolvimento de testes de DNA que detectam algumas das mutações que causam a doença Huntington e fibrose cística.

DIAGNÓSTICO MOLECULAR DE DOENÇAS HUMANAS

Genes mutantes que causam doenças humanas herdadas geralmente podem ser detectados por testes feitos no DNA genômico.

Uma vez que o gene responsável por uma doença humana tenha sido clonado e seqüenciado e a natureza das

mutações que causam o distúrbio seja conhecida, pode ser criado um teste molecular para os alelos mutantes. Esses testes podem ser feitos em pequenas quantidades de DNA usando-se PCR para amplificar o segmento de DNA de interesse (Fig. 15.11). Assim, eles podem ser feitos pré-natalmente em células fetais obtidas por amniocentese ou punção de vilosidades coriônicas ou mesmo em uma única célula de um pré-embrião produzida por fertilização *in vitro*.

Alguns diagnósticos moleculares envolvem testes simples quanto à presença ou à ausência de um sítio específico de clivagem de enzima de restrição no DNA. Por exemplo, a mutação que causa anemia falciforme (Cap. 13) remove um sítio de clivagem da enzima de restrição *MstII* (Fig. 17.6). O alelo Hb_{β}^S (falciforme) pode ser diferenciado do alelo normal de β -globina (Hb_{β}^A) amplificando-se parte do gene de β -globina por PCR, cortando o DNA amplificado com *MstII*, separando os fragmentos de restrição resultantes por eletroforese em gel de agarose, preparando uma transferência de Southern dos fragmentos separados e hibridizando o DNA na transferência com uma sonda que cobre o sítio da mutação. A sonda hibridizará com dois fragmentos pequenos do gene de β -globina normal, mas com apenas um fragmento do gene de β -globina falciforme (Fig. 17.6). Assim, este teste permite a detecção de heterozigotos bem como de homozigotos.

Alternativamente, podem ser sintetizados *primers* de PCR que são complementares a seqüências de DNA flanqueadoras da mutação falciforme no gene Hb_{β}^S e usados para amplificar este segmento do DNA genômico. O DNA amplificado pode ser tratado com *MstII* e os produtos da reação examinados por eletroforese em gel de agarose para verificar se os fragmentos amplificados foram ou não cortados pela *MstII*.

Para distúrbios herdados tais como doença Huntington e síndrome do X frágil, que resultam de regiões de repe-

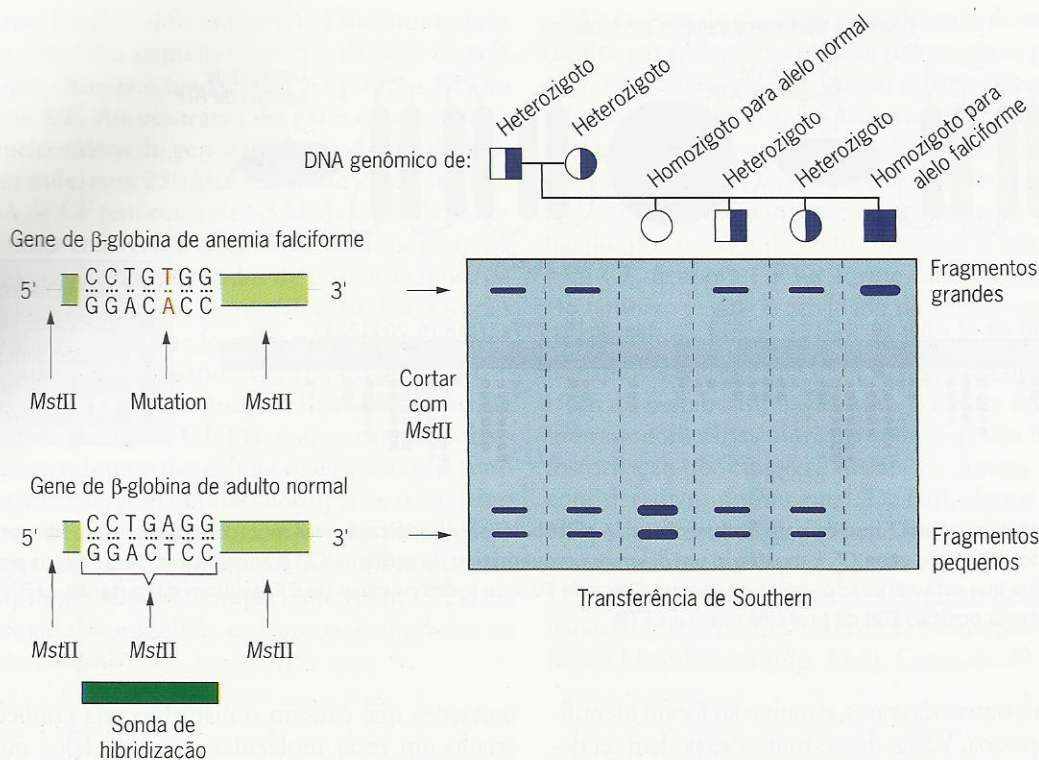


Fig. 17.6 ■ Detecção da mutação da hemoglobina falciforme por análise de transferência de Southern de DNA genômicos cortados com a enzima de restrição MstII.

tição expandida de trinucleotídeos em genes, podem ser usadas PCR e transferência de Southern para detectar alelos mutantes. O teste de DNA para o gene *huntingtina* é ilustrado na Fig. 17.2. Outros tipos de mutações podem ser detectados usando-se oligonucleotídeos aleloespecíficos como sondas em transferências de Southern genômicas. Este procedimento é ilustrado para a mutação $\Delta F508$ no gene *CF*, a causa mais freqüente de fibrose cística, em Enfoque Técnico: Detecção de um Gene Mutante Causador de Fibrose Cística, no Cap. 15. Uma vez caracterizadas as mutações responsáveis por uma doença, o desenvolvimento de testes de DNA para detectar as mais comuns geralmente é rotineiro. Claramente, a disponibilidade de testes diagnósticos para mutações que causam doenças humanas contribuiu muito para o campo da consulta genética, proporcionando valiosas informações para famílias nas quais ocorre o defeito genético.

■ PONTOS IMPORTANTES

- Genes mutantes que são responsáveis por distúrbios humanos herdados geralmente podem ser diagnosticados por testes de DNA.
- Os resultados de testes de DNA para genes mutantes que causam doenças herdas permitem que consultores genéticos informem as famílias dos riscos de terem filhos afetados.

TERAPIA GÊNICA HUMANA

Terapia gênica – a introdução de cópias funcionais de um gene em um indivíduo com duas cópias defeituosas do gene – é uma ferramenta potencial para o tratamento de doenças hereditárias humanas.



Das aproximadamente 5.000 doenças humanas herdadas catalogadas até hoje, apenas algumas são atualmente tratáveis. Para muitas destas doenças, o produto gênico ausente ou defeituoso não pode ser fornecido exogenamente, como insulina é dada a diabéticos. A maioria das enzimas são instáveis e não podem ser colocadas em sua forma funcional em locais de ação no corpo, pelo menos não em uma forma que forneça atividade a longo prazo. As membranas celulares são impermeáveis a moléculas grandes tais como proteínas. Assim, as enzimas têm que ser sintetizadas nas células onde elas são necessárias. Portanto, o tratamento de doenças herdadas é amplamente restrito a casos em que o metabólito ausente é uma pequena molécula que pode ser distribuída aos tecidos apropriados do corpo pelo sistema circulatório, ou os sintomas podem ser controlados por modificação da dieta da pessoa. Para muitas outras doenças herdadas, a **terapia gênica**

nica oferece o enfoque mais promissor para um tratamento bem-sucedido. A terapia gênica envolve a adição de uma cópia normal (tipo selvagem) de um gene ao genoma de uma pessoa portadora de cópias defeituosas do gene. Um gene que foi introduzido em uma célula ou organismo é chamado de **transgene** (*gene transferido*) para distingui-lo de genes endógenos, e o organismo que leva o gene introduzido é dito **transgênico**. Se a terapia gênica for bem-sucedida, o transgene irá sintetizar o produto gênico ausente e restaurará o fenótipo normal.

Antes de considerar exemplos específicos, precisamos discutir dois tipos de terapia gênica: **terapia gênica de células somáticas** ou **não herdável** e **terapia gênica herdável** ou **de linhagem germinativa**. Em animais superiores tais como humanos, as células reprodutivas ou da linhagem germinativa são produzidas por uma linhagem celular separada de todas as células da linhagem somática. Assim, a terapia gênica de células somáticas tratará os sintomas da doença da pessoa mas não irá curar a doença. Isto é, o(s) gene(s) defeituoso(s) ainda estará(ão) presente(s) nas células da linhagem germinativa do paciente após terapia de células somáticas e pode ser transmitido para sua prole. Todos os tratamentos envolvendo terapia gênica de doenças humanas que iremos discutir aqui são terapias gênicas de células somáticas. A terapia gênica de linhagem germinativa está sendo feita em camundongos e outros animais, mas não em humanos.

A distinção entre a terapia gênica de células somáticas e a da linhagem germinativa é importante quando discutirmos seres humanos. As preocupações geralmente manifestadas como "remendar a natureza" ou "brincar de Deus" aplicam-se a transferências de genes da linhagem germinativa, e não à terapia gênica de células somáticas. As principais considerações morais e éticas estão envolvidas em qualquer decisão sobre fazer modificações de genes humanos na linhagem germinativa. Em contraste, terapia gênica de células somáticas não é diferente de terapia enzimática (produto gênico) ou transplantes de células, tecidos e órgãos. Nos transplantes, órgãos inteiros, com todos os genes exógenos presentes no genoma de cada célula no órgão, são implantados em pacientes. Em terapias gênicas atuais de células somáticas, algumas das células do próprio paciente são removidas, reparadas e reimplantadas no paciente. Assim, a terapia gênica de células somáticas é menos complexa e menos ameaçadora à vida de uma pessoa do que um órgão transplantado.

Para realizar uma terapia gênica de célula somática, genes tipo selvagem devem ser introduzidos e expressos em células homozigotas ou hemizigotas* para um alelo mutante do gene. Em princípio, o gene tipo selvagem pode ser levado às células mutantes por qualquer um dentre vários procedimentos diferentes. Mais comumente, são usados vírus para levar o gene tipo selvagem para as células. No caso de vetores retrovirais, o transgene tipo selvagem é integrado, juntamente com o DNA retroviral, ao DNA da célula hospedeira. Assim, quando vetores retrovirais são usados, o transgene é

transmitido para todas as células da prole na linhagem celular afetada.

Com outros vetores virais, tais como os derivados de adenovírus, os transgenes estão presentes apenas transitoriamente em células hospedeiras, pois os genomas desses vírus replicam-se autonomamente e persistem apenas até que o sistema imunológico elimine os vírus junto com as células infectadas. A vantagem desses vetores sobre vetores retrovirais é que não são introduzidas mutações potencialmente prejudiciais durante a etapa de integração (Cap. 13). Entretanto, eles têm duas grandes desvantagens: (1) a expressão transgênica é transitória, durando apenas enquanto persistir a infecção viral, e (2) a maioria dos seres humanos exibe fortes respostas imunológicas a esses vírus, supostamente devido a exposição anterior ao mesmo vírus ou a um proximo relacionado. Por exemplo, nas primeiras tentativas de tratar fibrose cística por terapia gênica de células somáticas, um vetor adenoviral levando o gene *CF* foi inalado pelos pacientes, com a esperança de que células do pulmão fossem infectadas e produzissem suficiente quantidade do produto do gene *CF* para aliviar alguns dos sintomas da doença. Infelizmente, esses tratamentos não foram eficazes, pelo menos em parte devido a rápidas respostas imunológicas a tais vírus nas pessoas que receberam os tratamentos.

Com doenças tais como fibrose cística, em que terapia gênica eficaz irá precisar de expressão transgênica de longo prazo, os vetores adenovirais padrões provavelmente não funcionarão. Como a expressão transgênica é temporária, os tratamentos precisarão ser repetidos periodicamente. Entretanto, considerando que as respostas imunológicas secundárias são muito rápidas e eficientes (Cap. 23), os tratamentos subsequentes com o mesmo vetor viral provavelmente serão ineficazes.

A terapia gênica humana é feita sob orientações estritas desenvolvidas pelo National Institutes of Health (NIH). Cada procedimento proposto de terapia gênica é revisto por comitês tanto em nível local (instituição ou centro médico) quanto nacional (NIH). Vários requisitos devem ser atendidos antes que seja aprovado um procedimento de terapia gênica:

1. O gene deve ter sido clonado e bem caracterizado; isto é, deve estar disponível em forma pura.
2. Deve estar disponível um método eficiente para introduzir o gene no(s) tecido(s) ou nas células desejadas.
3. Os riscos da terapia gênica para o paciente devem ter sido cuidadosamente avaliados e demonstrados como mínimos.
4. A doença não deve ser tratável por outras estratégias.
5. Devem estar disponíveis dados de experimentos preliminares com modelos animais ou células humanas e deve ser indicado que a terapia gênica proposta deve ser eficaz.

Uma proposta de terapia gênica não será aprovada pelos comitês local e nacional até que eles estejam convencidos

*N.T.: Caso houvesse uma deleção de um alelo normal.

de que todas as condições citadas foram atendidas. Além disso, com a morte infeliz de Jesse Gelsinger em setembro de 1999, um rapaz com 18 anos de idade que tinha deficiência de ornitina-transcarbamilase, em consequência de uma intensa reação imunológica ao vetor adenoviral usado em sua terapia gênica experimental, os comitês de revisão estão sendo especialmente cautelosos em sua avaliação das propostas de terapia gênica.

O primeiro uso de terapia gênica em humanos ocorreu em 1990, quando uma menina de quatro anos de idade com **doença de imunodeficiência combinada grave por deficiência de adenosina-desaminase (ADA⁻SCID)** recebeu seu primeiro tratamento com um transgene. Vários outros pacientes com ADA⁻ SCID foram subsequentemente tratados, juntamente com algumas pessoas portadoras de outros distúrbios hereditários. SCID é uma rara doença autossômica do sistema imunológico. Pessoas com SCID essencialmente não têm sistema imunológico, e assim a menor infecção pode tornar-se grave e geralmente fatal. Alguns pacientes com SCID não têm uma enzima chamada adenosina-desaminase (ADA). Na ausência desta enzima, acumulam-se níveis tóxicos da forma fosforilada de seu substrato, desoxiadenosina, nos linfócitos T (glóbulos sangüíneos brancos essenciais para uma resposta imunológica) e as matam. Os linfócitos T estimulam células chamadas linfócitos B a desenvolver-se em plasmócitos produtores de anticorpos. Assim, na ausência de linfócitos T, não é possível uma resposta imunológica, e neonatos com ADA⁻ SCID raramente vivem mais que alguns anos.

Quatro fatores tornaram a ADA⁻ SCID um bom candidato para terapia gênica de células somáticas. Primeiro, o gene *ADA* foi um dos primeiros genes de doenças humanas a serem clonados e caracterizados. Segundo, leucócitos podem ser facilmente obtidos de pacientes com ADA⁻ SCID e reintroduzidos após serem adicionadas cópias funcionais do gene *ADA*. Terceiro, mesmo uma pequena quantidade de ADA funcional irá restaurar a função imunológica parcial. Quarto, a hiperprodução de ADA não parece ter efeitos tóxicos no paciente.

Como um pré-requisito para terapia gênica, um vetor de transferência gênica retroviral foi construído com o gene *ADA* humano sob o controle de um forte promotor viral (Fig. 17.7), e a expressão do gene quimérico foi demonstrada em leucócitos infectados com o vetor retroviral. O tratamento atual de ADA⁻ SCID por terapia gênica envolve quatro etapas sequenciais: (1) isolamento de leucócitos do paciente, (2) introdução de cópias funcionais do gene *ADA* nestas células, (3) demonstração de expressão do transgene nestas células e (4) infusão das células transgênicas no paciente (Fig. 17.8).

Embora o número de pacientes com ADA⁻ SCID que foram tratados por terapia gênica ainda seja pequeno, todos esses pacientes continuaram a receber tratamentos com ADA purificada, e a terapia resultou em melhora do funcionamento imunológico. A maioria está frequentando colégios normais, e alguns sobreviveram a doenças infecciosas tais como varíola. Esperava-se que os efeitos de terapia gênica usando leucócitos teria vida curta devido ao limitado tempo de vida destas células. Assim, infusões repetidas de leucócitos levando genes *ADA* funcionais têm sido feitas nos pacientes. O principal problema encontrado nas tentativas de tratar pacientes com ADA⁻ SCID por terapia gênica não tem sido o curto tempo de vida dos leucócitos, mas a expressão temporária dos transgenes *ADA* introduzidos. Dentro de poucas semanas após o tratamento, a transcrição do transgene *ADA* parou. A atividade dos promotores virais que controlam esta transcrição aparentemente cessou, indicando que é necessária pesquisa adicional para compreender por que transgenes geralmente são silenciados (expressão desligada) em seus novos hospedeiros.

Para evitar as limitações resultantes da vida curta dos leucócitos, as células tronco da medula óssea que originam os leucócitos (Cap. 23) podem ser usadas para tratar distúrbios imunológicos tais como ADA⁻ SCID. As células tronco modificadas devem produzir continuamente linfócitos T com o transgene *ADA* e podem fornecer um tratamento permanente ou de longo prazo da doença. Terapia gênica de

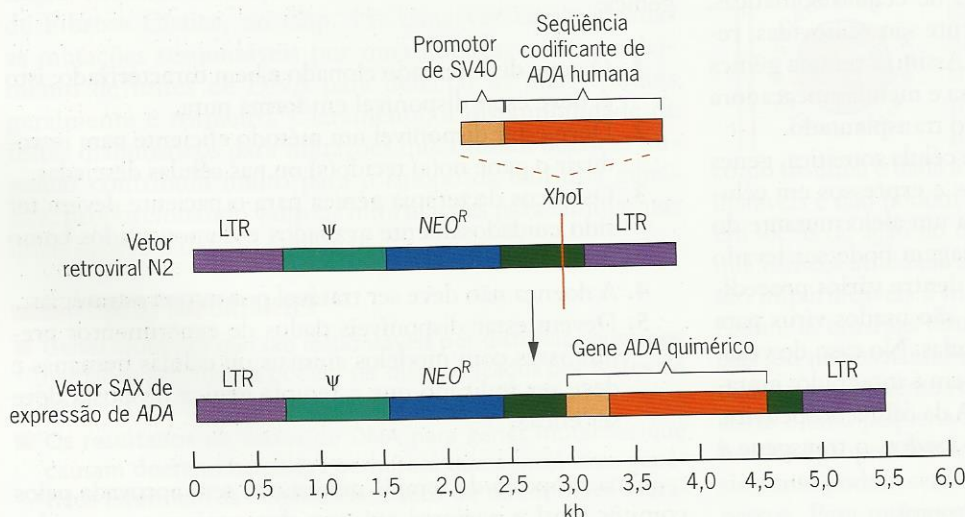


Fig. 17.7 ■ Construção do vetor SAX para transferência do gene de adenosina-desaminase humana (*ADA*) (para vírus 40 de símios, *ADA*, sítio *XhoI*). Neste vetor, a expressão da sequência codificante *ADA* humana é controlada por um promotor forte obtido do vírus 40 de símios. O gene *NEO^R* torna as células resistentes aos antibióticos neomicina, canamicina e G418. LTR refere-se às longas repetições terminais do vírus Moloney de leucemia murina; ψ é uma sequência de DNA necessária para embalar o vetor de DNA em vírions.

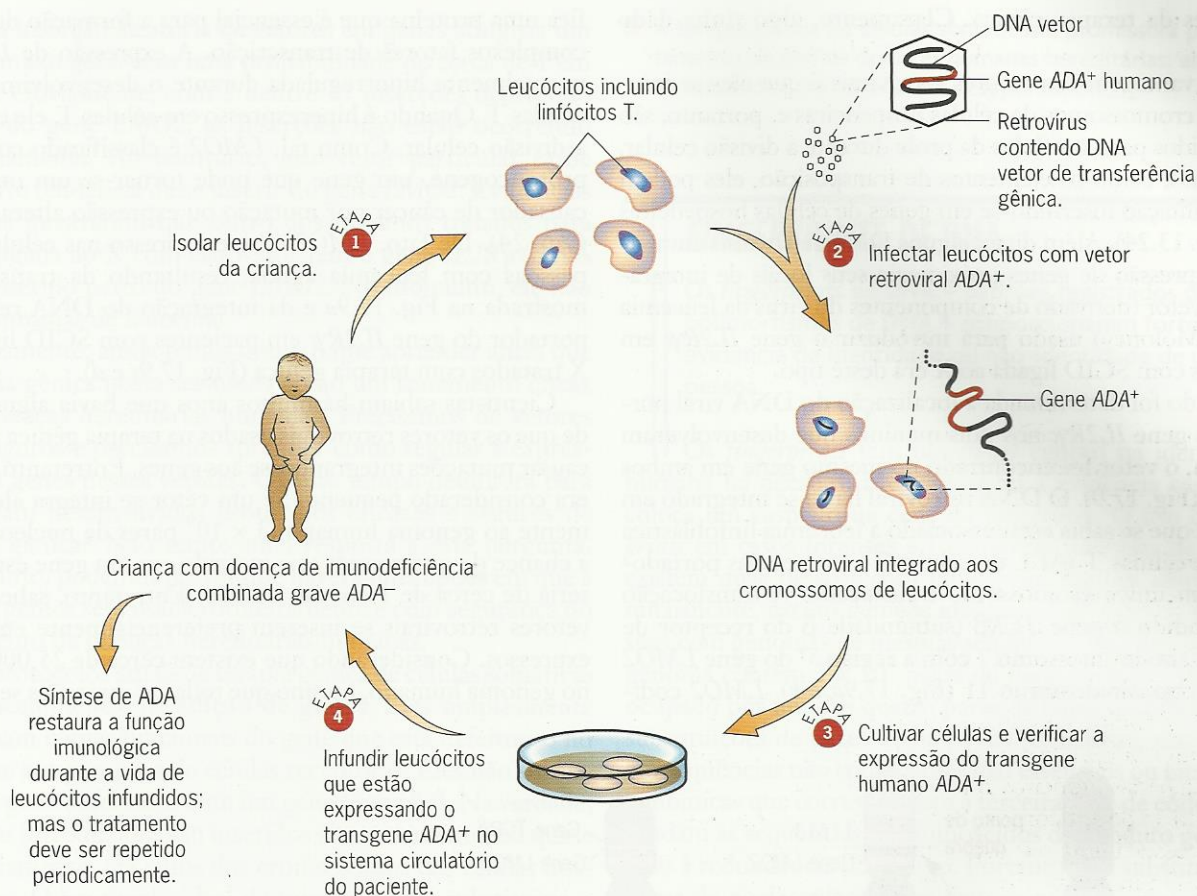


Fig. 17.8 ■ Tratamento gênico a partir de células somáticas da doença imunodeficiência combinada grave por deficiência de adenosina-desaminase (ADA⁻ SCID). A ADA⁻ SCID resulta de perda ou ausência de atividade da enzima adenosina-desaminase (ADA). A terapia gênica é feita isolando-se leucócitos do paciente, introduzindo uma cópia tipo selvagem do gene ADA nestas células com um vetor retroviral, verificando a expressão do transgene em células cultivadas e introduzindo os leucócitos transformados de volta no paciente.

células tronco foi primeiro usada para tratar duas crianças com ADA⁻ SCID em 1993, e este procedimento sem dúvida será o método de escolha no futuro.

Durante o ano 2000, médicos ingleses e franceses fizeram o que na época parecia ser o primeiro tratamento de terapia gênica de células somáticas bem-sucedido de pessoas com uma doença hereditária fatal. Eles trataram meninos com um tipo de SCID similar a ADA⁻ SCID já discutida mas causada por mutações em um gene no cromossomo X. Essa SCID ligada ao X resulta de perda ou inativação da subunidade γ do receptor de interleucina-2. Interleucina-2 é uma molécula sinalizadora necessária para o desenvolvimento de células do sistema imune. Entretanto, o polipeptídeo γ do receptor de interleucina-2 também é um componente de vários outros fatores de crescimento específicos de linfócitos. Coletivamente, eles estimulam o desenvolvimento de linfócitos B e T, células necessárias para a produção de plasmócitos produtores de anticorpos e células T *killer*, respectivamente (veja Cap. 23). Na ausência do polipeptídeo γ , uma pessoa não tem sistema imunológico funcional e raramente sobrevive por mais de alguns anos.

Como pessoas com ADA⁻ SCID, meninos com SCID ligada ao X parecem ser bons candidatos a tratamento por terapia gênica de células somáticas. Assim, o gene codificante da subunidade γ do receptor de interleucina-2 humano foi clonado, inserido em um vetor retroviral, introduzido em células tronco do sistema hematopoiético (precursoras de células do sistema circulatório) isoladas de pacientes com SCID ligada ao X e verificado quanto a expressão gênica enquanto as células ainda estavam no meio de cultura. Após verificar a expressão do gene (chamado de *IL2R γ* , de receptor γ de interleucina-2 comum), as células tronco foram transfundidas de volta para os pacientes com SCID dos quais haviam sido isoladas. Durante os dois anos seguintes, 14 meninos com SCID ligada ao X foram tratados. Em todos os 14 casos, a terapia gênica curou a imunodeficiência, resultando em níveis normais de células T dentro de alguns meses após o tratamento. Assim, por dois anos, tudo indicava que a terapia gênica havia sido um grande sucesso. Então um dos meninos desenvolveu uma leucemia de células T aguda, um câncer dos leucócitos. Posteriormente, a mesma leucemia de células T foi detectada em outro dos

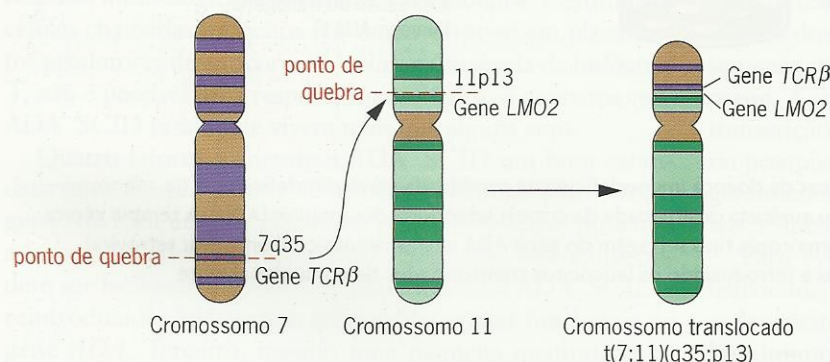
pacientes da terapia gênica. Claramente, algo tinha dado errado.

Uma vantagem dos vetores retrovirais é que eles se inserem nos cromossomos de células hospedeiras e, portanto, são transmitidos para as células da prole durante a divisão celular. Entretanto, como os elementos de transposição, eles podem causar mutação inserindo-se em genes de células hospedeiras (veja Fig. 13.24). Além disso, alguns DNA retrovirais aumentam a expressão de genes próximos a seus locais de integração, e o vetor (derivado de componentes do vírus da leucemia murina Moloney) usado para introduzir o gene *IL2R γ* em pacientes com SCID ligada ao X era deste tipo.

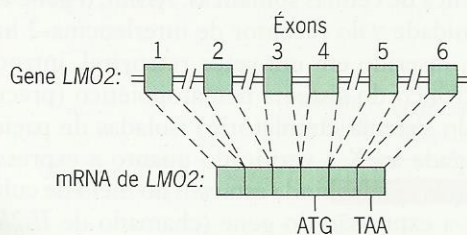
Quando foi determinada a localização do DNA viral portador do gene *IL2R γ* nos dois meninos que desenvolveram leucemia, o vetor foi encontrado no mesmo gene em ambos os casos (Fig. 17.9). O DNA retroviral havia se integrado em um gene que se sabia estar associado a leucemia linfoblástica aguda de células T (ALL de células T) em pessoas portadoras de um único cromossomo translocado. A translocação havia fundido o gene *TCR β* (subunidade β do receptor de células T) no cromossomo 7 com a região 5' do gene *LMO2* (só LIM) no cromossomo 11 (Fig. 17.9a e b). *LMO2* codi-

fica uma proteína que é essencial para a formação de alguns complexos fatores de transcrição. A expressão de *LMO2* é normalmente hiporregulada durante o desenvolvimento de células T. Quando é hiperexpresso em células T, ele estimula a divisão celular. Como tal, *LMO2* é classificado como um protooncogene, um gene que pode tornar-se um oncogene causador de câncer por mutação ou expressão alterada (veja Cap. 24). De fato, *LMO2* é hiperexpresso nas células T de pessoas com leucemia aguda, resultando da translocação mostrada na Fig. 17.9a e da integração do DNA retroviral portador do gene *IL2R γ* em pacientes com SCID ligada ao X tratados com terapia gênica (Fig. 17.9c e d).

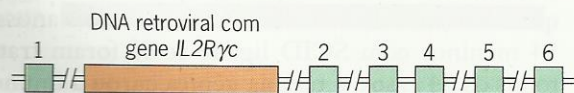
Cientistas sabiam há muitos anos que havia algum risco de que os vetores retrovirais usados na terapia gênica podiam causar mutações integrando-se aos genes. Entretanto, o risco era considerado pequeno. Se um vetor se integra aleatoriamente ao genoma humano (3×10^9 pares de nucleotídeos), a chance de que o vetor seja inserido em um gene específico seria de cerca de 1 em um milhão. Entretanto, sabe-se que vetores retrovirais se inserem preferencialmente em genes expressos. Considerando que existem cerca de 25.000 genes no genoma humano, mesmo que todas as inserções sejam em



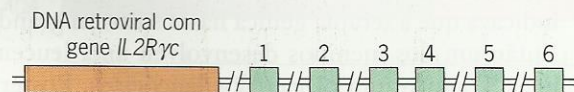
(a) Identificação do oncogene *LMO2* por sua associação com uma translocação entre os cromossomos 7 e 11 em pessoas com leucemia linfoblástica aguda de células T (ALL de células T)



(b) Estrutura do gene normal *LMO2* e mRNA.



(c) Sítio de inserção de DNA retroviral no menino 1 com leucemia.



(d) Sítio de inserção de DNA retroviral no menino 2 com leucemia.

Fig. 17.9 ■ O gene *LMO2* (gene 2 LIM) codifica uma pequena proteína (156 aminoácidos) que funciona como uma ponte unindo diferentes fatores de transcrição. Ele foi identificado em estudos de pessoas com leucemia linfoblástica aguda de células T (ALL de células T, um câncer que afeta leucócitos). (a) Nesses pacientes, ocorreu uma translocação entre os cromossomos 7 e 11. Essa translocação moveu o gene *TCR β* (subunidade β do receptor de células T) no cromossomo 7 para perto do gene *LMO2* no cromossomo 11 e resultou na hiperexpressão de *LMO2*. Quando hiperexpresso, *LMO2* comporta-se como um oncogene (gene causador de câncer; veja Cap. 24) em uma via que resulta em leucemia de células T. (b) As estruturas do gene *LMO2* tipo selvagem e do mRNA. O gene contém seis éxons, com os éxons 4–6 codificando a proteína *LMO2*. São mostradas as posições dos sítios de início da tradução (ATG) e de término (TAA). Nos dois pacientes submetidos a terapia gênica que desenvolveram leucemia, o vetor retroviral portador do gene *IL2R γ* (subunidade comum codificante do receptor gama de interleucina-2) integrou (c) os éxons 1 e 2 (menino 1) ou (d) o éxon 1 antecessor (menino 2). Em ambos os casos, o DNA retroviral integrado aumentou a expressão do gene *LMO2* nos leucócitos, onde sua expressão normalmente é diminuída.

genes, a inserção aleatória de vetores em genes atingiria um determinado gene com uma probabilidade de cerca de 1 em 25.000. Obviamente, com 2 dentre 15 inserções ocorrendo dentro do gene *LMO2*, as inserções não estão ocorrendo aleatoriamente. Ao contrário, este vetor particular exibe uma forte tendência de se inserir no gene *LMO2*. Evidências recentes mostraram que outro dos meninos tratados para SCID ligada ao X com este vetor tem o gene *IL2R γ* /DNA vetor inserido perto do gene *LMO2*. No momento, ele não exibe sintomas de leucemia.

Claramente, ainda temos muito o que aprender antes que a terapia gênica possa ser usada como um tratamento eficaz para doenças hereditárias humanas. Precisamos de vetores mais seguros e precisamos aprender como regular a expressão dos genes nesses vetores. Quanto tempo demorará para que sejam desenvolvidos protocolos eficazes e seguros de terapia gênica? Não temos uma resposta a esta pergunta. Entretanto, podemos prever que haverá uma época em que a terapia gênica será usada rotineiramente e com segurança no tratamento de doenças hereditárias humanas.

Os protocolos atuais de terapia gênica de células somáticas são procedimentos de **adição de genes**. Eles simplesmente adicionam cópias funcionais do gene que está defeituoso no paciente aos genomas de células receptoras. Eles não substituem o gene defeituoso com um gene funcional. Na verdade, os genes introduzidos são inseridos aleatoriamente, ou quase aleatoriamente, em sítios dos cromossomos das células hospedeiras. O protocolo ideal de terapia gênica substituiria o gene defeituoso por um gene funcional. As **substituições de genes** seriam mediadas por recombinação homóloga e colocariam o gene introduzido em seu local normal no genoma hospedeiro. Em humanos, substituições gênicas são geralmente referidas como *transferências gênicas direcionadas*. Oliver Smithies e colaboradores primeiro usaram recombinação homóloga para direcionar seqüências de DNA para o locus de β -globina de células humanas de culturas de tecidos em 1985. Entretanto, a freqüência da transferência do gene direcionado era muito baixa (cerca de 10^{-5}). Desde então, Smithies, Mario Capecchi e outros desenvolveram vetores de direcionamento gênico melhorados e estratégias de seleção. Como resultado, são possíveis substituições gênicas direcionadas mais eficientes, e células com a substituição gênica desejada podem ser mais facilmente identificadas. No futuro, substituições gênicas direcionadas provavelmente serão o método de escolha para tratamento por terapia gênica de células somáticas de doenças humanas.

PONTOS IMPORTANTES

- A terapia gênica envolve a adição de uma cópia normal (tipo selvagem) de um gene ao genoma de uma pessoa que possui cópias defeituosas do gene.
- Embora a terapia gênica de células somáticas restaure efetivamente o funcionamento imunológico em meninos com doença de imunodeficiência combinada grave ligada ao X, três dos meninos subsequentemente desenvolveram leucemia ou distúrbios do tipo leucemia.

- A terapia gênica de células somáticas é promissora para o tratamento de muitas doenças humanas hereditárias; entretanto, os resultados até hoje foram desapontadores.

FINGERPRINTS DE DNA

Os *fingerprints* de DNA – padrões registrados de polimorfismos de DNA – proporcionaram forte evidência de identidade ou não-identidade de uma pessoa.

Os *fingerprints* têm um papel central na identificação humana há décadas. Eles fornecem a evidência central que coloca um suspeito na cena de um crime. O uso de *fingerprints* em casos forenses é baseado na premissa de que não existem duas pessoas com impressões digitais idênticas. Similarmente, exceto gêmeos idênticos, não existem duas pessoas com genomas de mesmas seqüências de nucleotídeos. O genoma contém 3×10^9 pares de nucleotídeos. Cada sítio é ocupado por um dos quatro pares de bases no DNA. Muitas substituições de pares de bases são silenciosas; elas ocorrem em seqüências não codificantes não essenciais ou em posições genômicas que correspondem à terceira base de códons e não mudam as seqüências de aminoácidos do produto gênico devido à redundância do código. Portanto, tais substituições de pares de nucleotídeos acumulam-se nos genomas durante o curso da evolução. Além disso, duplicações e deleções de seqüências de DNA e outros rearranjos genômicos contribuem para a divergência evolutiva dos genomas. Evidências recentes demonstraram que o genoma humano contém grandes famílias de polimorfismos de DNA de muitos tipos diferentes, polimorfismos que podem fornecer valiosas evidências em casos de identidade incerta. Esses polimorfismos podem ser usados para produzir **fingerprints de DNA**, padrões específicos de bandamento em transferências de Southern de DNA genômico clivado com uma enzima de restrição específica e hibridizado com sondas apropriadas de DNA.

O poder e a utilidade do procedimento de *fingerprinting* de DNA em casos de identificação de pessoas são óbvios para qualquer um que conheça genética molecular e as técnicas usadas na produção de impressões de DNA. As controvérsias quanto ao uso dos *fingerprints* de DNA em casos forenses estão relacionadas com a competência dos laboratórios de pesquisa envolvidos, com a probabilidade de erro humano em produzir as impressões e com os métodos para calcular a probabilidade de que duas pessoas tenham *fingerprints* idênticos. Para fazer estimativas precisas da probabilidade de impressões idênticas, pesquisadores devem ter informações confiáveis sobre a freqüência dos polimorfismos na população em questão. Por exemplo, se a endogamia (reproduções entre indivíduos aparentados) é comum na população, a probabilidade de *fingerprints* idênticos irá aumentar. Assim, estimativas precisas da probabilidade de que dois indivíduos tenham *fingerprints* correspondentes requerem informações confiáveis sobre as

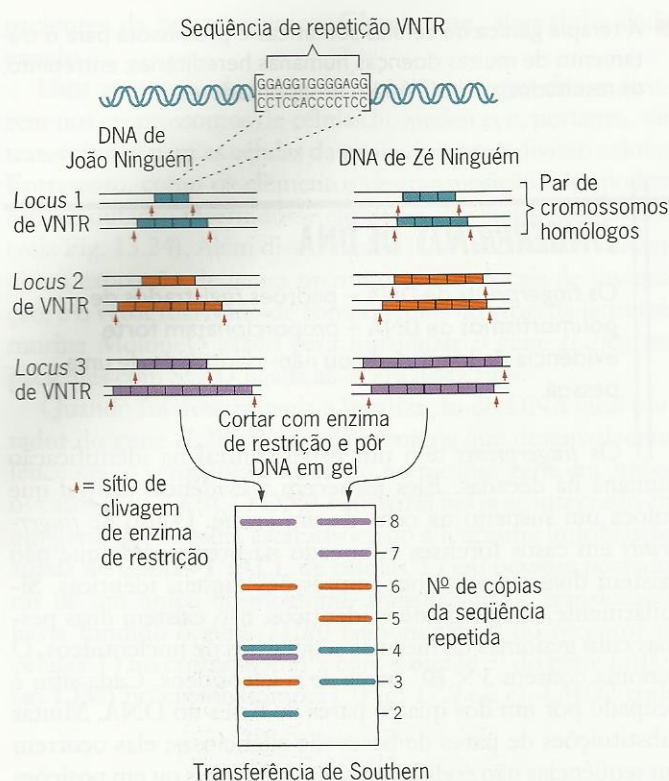


Fig. 17.10 ■ Diagrama simplificado do uso de número variável de repetições em tandem na preparação de *fingerprints* de DNA.

freqüências dos polimorfismos na população relevante. Dados obtidos de uma população nunca devem ser extrapolados para outra população, pois freqüências diferentes de polimorfismos podem estar presentes em populações diferentes.

O *fingerprint* de DNA fornece uma poderosa ferramenta forense se for apropriadamente usado. *Fingerprints* de DNA podem ser preparados a partir de quantidades pequenas de sangue, sêmen, bulbos capilares ou outras células. O DNA é extraído destas células, amplificado por PCR e analisado com sondas de DNA cuidadosamente escolhidas pelo procedimento de transferência de Southern (Cap. 15). *Fingerprint* às vezes pode ser feito com DNA isolado de tecidos preservados de pessoas muito depois de sua morte. Como previamente mencionado, o genoma humano contém várias seqüências curtas de DNA que estão presentes como repetições em tandem de tamanhos variados em muitos locais cromossômicos. Esses números variáveis de repetições em tandem, ou VNTR, são componentes importantes dos *fingerprints* de DNA (Fig. 17.10). Embora aplicáveis em todos os casos de identidade questionável, eles se mostraram especialmente úteis em casos forenses e de paternidade.

Testes de Paternidade

No passado, casos de incerteza de paternidade geralmente eram decididos comparando-se os tipos sanguíneos da

criança, da mãe e dos possíveis pais. Dados de grupos sanguíneos podem ser usados para provar que homens com determinados grupos sanguíneos não poderiam ser o pai da criança. Infelizmente, tais comparações de tipo sanguíneo contribuem pouco para uma identificação positiva do pai. Em contraste, os *fingerprints* de DNA não só excluem pais identificados erradamente como também chegam perto de uma identificação positiva do verdadeiro pai. Amostras de DNA são obtidas de células da criança, da mãe e dos possíveis pais, e os *fingerprints* de DNA são preparados como descrito na Fig. 17.10. Quando são comparados os *fingerprints*, todas as bandas na impressão do DNA da criança devem estar presentes nas impressões de DNA combinadas dos genitores. Para cada par de cromossomos homólogos, a criança tem que ter recebido um de cada genitor. Assim, aproximadamente metade das bandas na impressão do DNA da criança resultará de seqüências de DNA herdadas da mãe e metade de seqüências de DNA herdadas do pai.

A Fig. 17.11 mostra os *fingerprints* de DNA de uma criança, da mãe e de dois homens suspeitos de serem o pai da criança. Neste caso, as impressões de DNA indicam que o segundo suspeito é provavelmente o pai biológico da criança. A precisão dos *fingerprints* de DNA na identificação das relações pai-filho pode ser acentuada aumentando-se o número de sondas de hibridização usadas na análise. Com o uso de mais sondas, mais polimorfismos podem ser examinados e uma proporção maior dos genomas da criança e dos genitores pode ser comparada. O resultado é uma identificação mais confiável.

Aplicações Forenses

Fingerprints de DNA foram primeiro usados como evidência em um caso criminal em 1988. Em 1987, um juiz na Flórida negou a solicitação do promotor de apresentar interpretações estatísticas de evidências de DNA contra um acusado de estupro. Após um julgamento errado, o suspeito foi liberado. Três meses depois, ele estava novamente no tribunal, acusado de outro estupro. Desta vez, o juiz permitiu que o promotor apresentasse uma análise estatística dos dados, com base em levantamentos populacionais apropriados. A análise mostrou que o *fingerprint* de DNA preparado a partir de sêmen recuperado da vítima tinha uma probabilidade de cerca de um em 10 bilhões de que a correspondência do *fingerprint* do DNA do suspeito fosse apenas por acaso. Desta vez, o suspeito foi condenado. Não há dúvida quanto ao valor do teste de DNA em casos forenses deste tipo quando uma boa amostra de células ou tecidos é obtida da cena do crime. Se forem feitos cuidadosamente por cientistas treinados e interpretados com base em dados das distribuições populacionais dos polimorfismos envolvidos, os *fingerprints* de DNA podem ser uma ferramenta muito necessária e poderosa para a contínua luta contra o crime.

A Fig. 17.12 ilustra um tipo de *fingerprint* de DNA usado em casos forenses, os VNTR. O *fingerprint* de DNA pre-

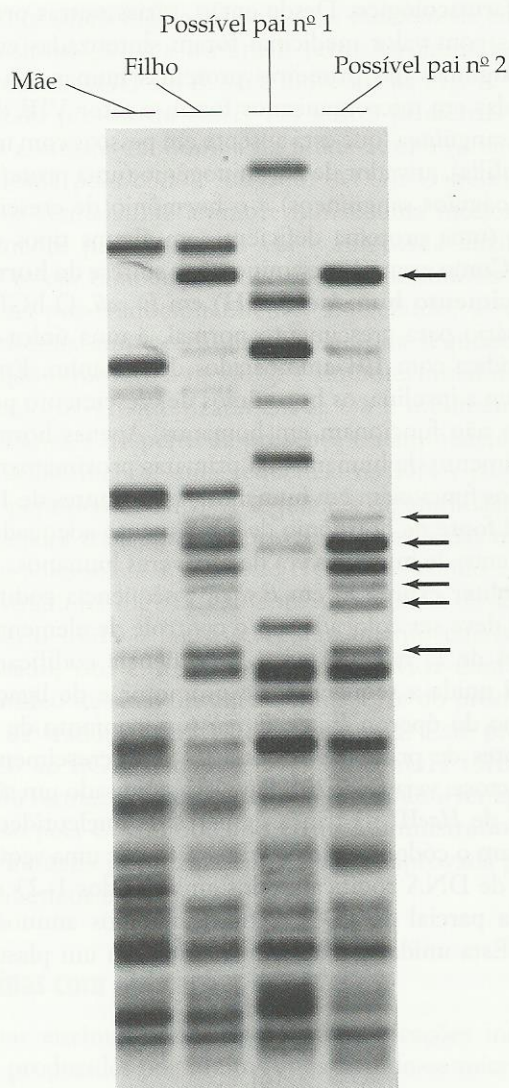


Fig. 17.11 ■ Fingerprints de DNA de uma mãe, seu filho e dois homens, cada um dos quais diz ser o pai da criança. As setas marcam as bandas que identificam o homem nº 2 como o pai biológico.

parado com uma mancha de sangue de uma cena de crime corresponde às impressões de DNA do suspeito 1, mas não às dos outros dois suspeitos. Logicamente, essa correspondência de DNA em si não prova que o suspeito 1 tenha cometido o crime, mas, se combinada com impressões de DNA adicionais e evidências de apoio, pode dar fortes indicações de que o suspeito 1 tenha estado na cena do crime. Talvez, mais importante, essas impressões mostrem claramente que as células sanguíneas na mancha não sejam dos outros dois suspeitos. Assim, os *fingerprints* de DNA são valiosos em reduzir a frequência de condenações erradas.

Combinando os *fingerprints* de VNTR produzidos com manchas usando-se outros tipos de sondas de DNA, a possibilidade de que os *fingerprints* de DNA de duas pessoas correspondam apenas por acaso pode ser praticamente eli-

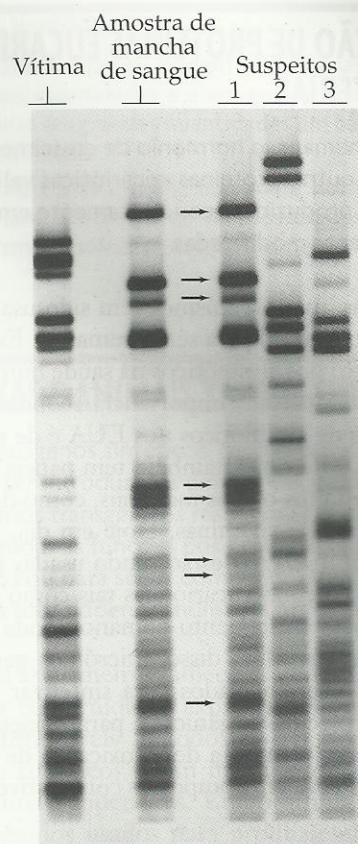


Fig. 17.12 ■ Fingerprints de DNA preparados de DNA isolado de uma mancha de sangue no local de um crime e de sangue obtido de três pessoas suspeitas de cometer o crime. As setas indicam fragmentos de DNA do suspeito 1 – não presentes nos suspeitos 2 ou 3 – que correspondem à mancha de sangue da cena do crime.

minada. A lógica subjacente ao uso de *fingerprints* de DNA para identificar pessoas é que o DNA de cada pessoa tem uma sequência única de nucleotídeos. A despeito da explosão da população humana, existem muito mais combinações possíveis dos quatro pares de bases em um genoma humano de 3×10^9 pares de bases do que humanos em nosso planeta. Assim, exceto no caso de gêmeos idênticos*, não existem dois seres humanos que tenham genomas idênticos. Os *fingerprints* de DNA fornecem uma ferramenta por meio da qual estas diferenças podem ser detectadas e registradas, como há décadas são registradas as impressões digitais.

PONTOS IMPORTANTES

- Os *fingerprints* de DNA detectam e registram polimorfismos nos genomas de indivíduos.
- As impressões do DNA fornecem forte evidência da identidade individual, evidência que pode ser extremamente valiosa em casos forenses e de paternidade.

*N.T.: Na verdade, nem mesmo os gêmeos monozigóticos são 100% idênticos. Você imagina por quê?

PRODUÇÃO DE PROTEÍNAS EUCARIÓTICAS EM BACTÉRIAS

A insulina humana, o hormônio de crescimento humano, e outras proteínas eucarióticas valiosas podem ser produzidas economicamente em bactérias geneticamente modificadas.

Por décadas, microrganismos têm sido usados para gerar produtos importantes para seres humanos. Estamos todos cientes do impacto dos antibióticos na saúde humana. Poucos de nós estão cientes de sua importância econômica. O valor total de mercado de antibióticos nos EUA é de mais de US\$ 2 bilhões anuais. Micróbios também têm papéis importantes na produção de muitos outros materiais, como drogas contra fungos, aminoácidos e vitaminas. Hoje em dia, graças à engenharia genética, bactérias estão sendo usadas na produção de importantes proteínas eucarióticas tais como insulina humana, hormônio de crescimento humano e toda a família de interferons humanos. Além disso, micróbios geneticamente modificados estão sendo usados para sintetizar valiosas enzimas e outras moléculas orgânicas e para fornecer uma maquinaria metabólica visando a desintoxicação de poluentes e conversão de biomassa em compostos combustíveis.

Hormônio de Crescimento Humano

Em 1982, a insulina humana tornou-se o primeiro sucesso comercial das novas tecnologias de DNA recombinante no

campo farmacológico. Desde então, várias outras proteínas humanas com valor medicinal foram sintetizadas em bactérias. Algumas das primeiras proteínas humanas a serem produzidas em microrganismos foram o fator VIII de coagulação sanguínea (que está ausente em pessoas com um tipo de hemofilia), ativador de plasminogênio (uma proteína que desfaz coágulos sanguíneos) e o hormônio de crescimento humano (uma proteína deficiente em alguns tipos de nanismo). Como exemplo, examinemos a síntese do **hormônio de crescimento humano (hGH)** em *E. coli*. O hGH, que é necessário para crescimento normal, é uma única cadeia polipeptídica com 191 aminoácidos de tamanho. Em contraste com a insulina, os hormônios de crescimento porcino e bovino não funcionam em humanos. Apenas hormônios de crescimento de humanos ou primatas proximamente relacionados funcionam em humanos. Assim, antes de 1985, a principal fonte de hormônio de crescimento adequado para o tratamento de humanos era de cadáveres humanos.

Para obter expressão em *E. coli*, a sequência codificante de hGH deve ser colocada sob o controle de elementos reguladores de *E. coli*. Portanto, a sequência codificante de hGH foi unida a sequências de promotor e de ligação ao ribossomo do óperon *lac* de *E. coli* (um conjunto de genes codificantes de proteínas necessárias para crescimento no açúcar lactose; veja Cap. 20). Para isto, foi usado um sítio de clivagem de *HaeIII* na trinca de pares de nucleotídeos que especificam o códon 24 do hGH para fundir uma sequência sintética de DNA codificante dos aminoácidos 1–23 a uma sequência parcial de cDNA codificante dos aminoácidos 24–191. Esta unidade foi então inserida em um plasmídeo

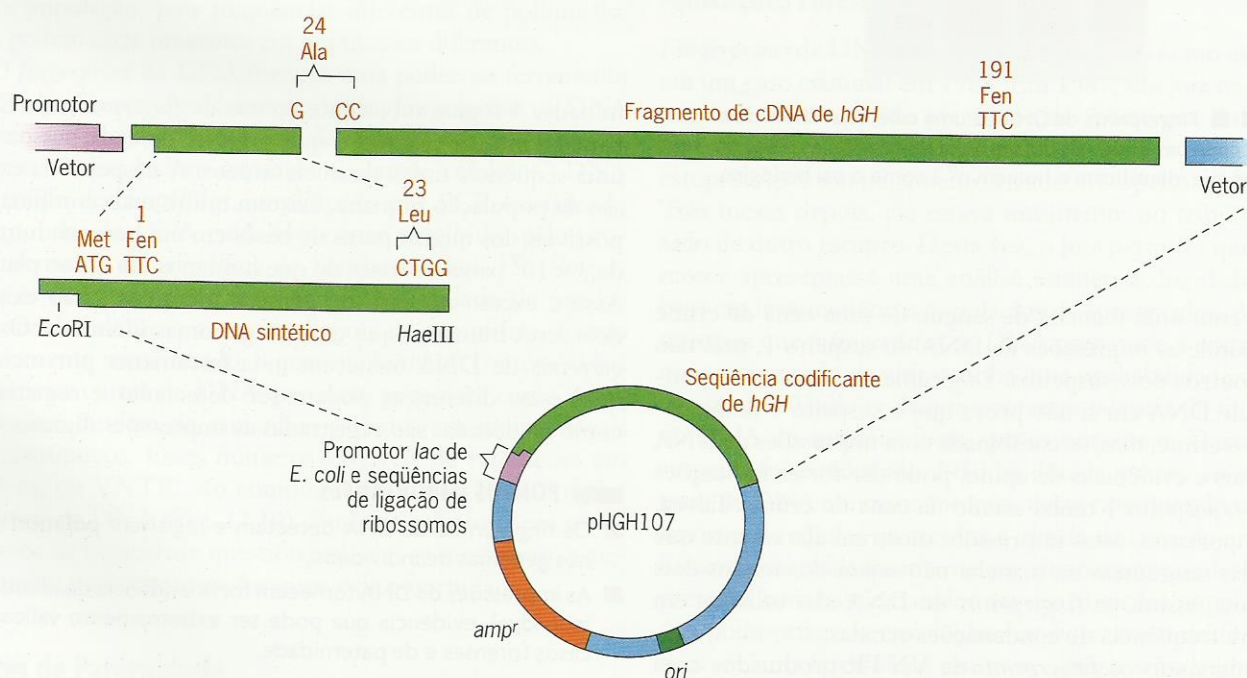


Fig. 17.13 ■ Estrutura do primeiro vetor usado para produzir o hormônio de crescimento humano (hGH) em *E. coli*. O gene *amp^r* fornece resistência a ampicilina; *ori* é a origem de replicação do plasmídeo. Os aminoácidos são numerados de um a 191 começando no terminal amino.

que leva os sinais reguladores de *lac* e introduzida em *E. coli* por transformação. A estrutura do primeiro plasmídeo usado para produzir hGH em *E. coli* é mostrada na Fig. 17.13.

O hGH produzido em *E. coli* nestes primeiros experimentos continha metionina na ponta amino (a metionina especificada pelo códon iniciador ATG). O hGH nativo tem uma fenilalanina aminoterminal; uma metionina está inicialmente presente, mas ela é removida enzimaticamente. *E. coli* também remove muitas metioninas aminoterminalis pós-traducionalmente. Entretanto, a remoção da metionina aminoterminal depende da sequência, e as *E. coli* não removem a metionina aminoterminal do hGH. Todavia, o hGH sintetizado em *E. coli* é totalmente ativo em humanos a despeito da presença do aminoácido extra. Mais recentemente, uma sequência de DNA codificante de um peptídeo de sinal (a sequência de aminoácidos necessária para transporte de proteínas através das membranas) foi adicionada a um gene *HGH* construído e similar ao mostrado na Fig. 17.13. Com a sequência de sinal adicionada, o hGH é tanto secretado quanto corretamente processado; isto é, a metionina é removida com o resto do peptídeo de sinal durante o transporte do produto primário de tradução através da membrana. Este produto é idêntico ao hGH nativo. Em 1985, o hGH tornou-se o segundo fármaco geneticamente construído a ser aprovado para uso pela U. S. Food and Drug Administration. A insulina humana produzida em *E. coli* foi aprovada para uso por diabéticos em 1982.

Proteínas com Aplicações Industriais

Algumas enzimas com importantes aplicações industriais foram produzidas por muitos anos usando-se microrganismos para fazer a sua síntese. Por exemplo, proteases foram produzidas pelo *Bacillus licheniformis* e outras bactérias. Tais proteases foram amplamente empregadas como auxiliares em detergentes e em menores quantidades como amaciantes de carnes e auxiliares digestivos em rações animais. Amilases têm sido amplamente usadas para degradar carboidratos complexos tais como amido em glicose. A glicose é então convertida em frutose com a enzima glicose-isomerase, e essa frutose é usada como adoçante de alimentos. As amilases e a glicose-isomerase são todas feitas por processos microbiológicos.

A proteína renina é usada para fazer queijos. Antes do advento da engenharia genética, a renina era extraída do quarto estômago de bovinos. Bactérias geneticamente modificadas hoje são usadas para a produção comercial de renina. Tais exemplos são todos de proteínas que tiveram importantes aplicações industriais por algum tempo. No futuro, podemos esperar que muitas enzimas adicionais sejam produzidas e usadas em aplicações industriais em razão da facilidade de produzir tais proteínas por meio de microrganismos recombinantes (ou por plantas e animais transgênicos; veja a seção seguinte).

PONTOS IMPORTANTES

- Proteínas valiosas que só podem ser isoladas de eucariontes em pequenas quantidades e a um grande custo agora podem ser produzidas em grandes quantidades em bactérias geneticamente modificadas.
- Proteínas tais como insulina humana e hormônio de crescimento humano são fármacos valiosos usados para tratar diabetes e nanismo hipofisário, respectivamente.

PLANTAS E ANIMAIS TRANSGÊNICOS

Genes exógenos sintéticos, modificados ou outros podem ser introduzidos em plantas e animais, e os organismos transgênicos resultantes podem ser usados para estudar as funções dos genes, para produzir novos produtos ou para servir como modelos animais de estudos sobre doenças humanas hereditárias.

Embora as primeiras moléculas de DNA recombinante tenham sido construídas e expressas em microrganismos, hoje estão disponíveis métodos que permitem que genes exógenos ou sintéticos sejam introduzidos e expressos em plantas e animais superiores. Embora uma discussão completa dos métodos usados para produzir plantas e animais transgênicos esteja além do escopo deste livro, vamos discutir alguns dos procedimentos comumente usados e algumas das aplicações iniciais de tecnologias de DNA recombinante em cruzamentos de plantas e animais.

Animais Transgênicos: Microinjeção de DNA em Zigotos

Dois métodos são predominantemente usados para produzir animais transgênicos: (1) microinjeção de DNA em zigotos e (2) infecção de embriões pré-implantação com vetores retrovirais. Só discutiremos aqui o procedimento de microinjeção. Veja na Fig. 17.8 uma ilustração do procedimento retroviral.

A maioria dos animais transgênicos estudados hoje em dia foram produzidos por microinjeção de DNA em zigotos. Antes da microinjeção, os ovócitos são removidos cirurgicamente do genitor feminino e são fertilizados *in vitro*. O DNA é então microinjetado no pró-núcleo masculino (o núcleo haplóide contribuído pelo espermatozoide, antes da fusão nuclear) do ovócito fertilizado por meio de uma agulha de vidro com ponta muito fina (Fig. 17.14). Geralmente, centenas a milhares de cópias do gene de interesse são injetadas em cada ovócito fertilizado, e em geral podem ocorrer várias integrações. Surpreendentemente, quando várias cópias se integram ao genoma, em geral o fazem em tandem, em um único sítio cromossômico. A integração de moléculas de DNA injetadas parece ocorrer em sítios aleatórios no genoma.

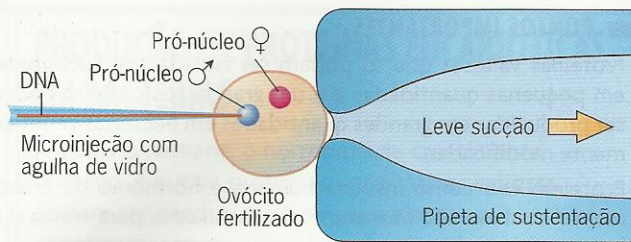


Fig. 17.14 ■ A produção de animais transgênicos por microinjeção de DNA em ovócitos fertilizados.

Como o DNA é injetado no ovócito fertilizado, a integração das moléculas de DNA injetadas geralmente ocorre cedo durante o desenvolvimento embrionário. Como resultado, algumas células da linhagem germinativa podem levar o transgene. Como seria esperado, os animais que se desenvolvem de ovócitos injetados, referidos como a **geração G_0** , são quase sempre mosaicos genéticos, com algumas células somáticas levando o transgene e outras não. Os animais transgênicos iniciais (G_0) devem ser cruzados e a prole G_1 produzida para se obter animais em que todas as células levam o transgene. Na maioria dos casos em que sua herança foi estudada, os transgenes foram transmitidos para a prole de modo estável. Em alguns animais tais como o nematódeo do solo *Caenorhabditis elegans*, o transgene injetado não é integrado de modo estável ao genoma, mas é transmitido para a prole como uma molécula de DNA de replicação autônoma como os plasmídeos de bactérias.

Camundongos transgênicos são produzidos rotineiramente em laboratórios em todo o mundo. Eles são ferramentas valiosas para se estudar a expressão gênica em mamíferos e um excelente sistema modelo com o qual testar vários vetores de transferência de genes e metodologias para possível uso em humanos. Centenas de transgenes diferentes foram estudados em camundongos, e, na maioria dos casos, os transgenes apresentam padrões normais de herança, indicando que eles foram integrados ao genoma hospedeiro. Discutiremos algumas das aplicações de camundongos transgênicos em estudos de desenvolvimento no Cap. 22.

Um dos primeiros experimentos com camundongos transgênicos mostrou que a taxa de crescimento podia ser aumentada quando genes de hormônio de crescimento de rato, bovinos ou seres humanos eram expressos em camundongos (Fig. 17.15). Isto levou criadores a indagar se a introdução de (1) cópias extras do gene homólogo (mesma espécie) de hormônio de crescimento ou (2) cópias dos genes heterólogos do hormônio de crescimento de espécies correlatas poderiam resultar em animais domésticos com taxas de crescimento acentuadas. Foram produzidos porcos transgênicos com a esperança de que níveis aumentados de hormônio de crescimento pudessem produzir porcos mais magros com aumento de qualidade da carne e crescimento mais rápido. Outros cientistas introduziram transgenes de hormônio de crescimento em peixes e galinhas com objetivos similares.



Fig. 17.15 ■ O camundongo transgênico da esquerda, que leva um gene quimérico de hormônio de crescimento humano, tem cerca do dobro do tamanho do camundongo controle à direita.

Os experimentos com porcos transgênicos indicaram que as taxas de crescimento foram aumentadas, não em dietas padrões, mas apenas em dietas ricas em proteína. Entretanto, os porcos transgênicos apresentaram-se mais magros que os controles, como esperado, porque o hormônio de crescimento favorece a síntese de proteínas, e não a de gorduras. Infelizmente, os porcos transgênicos também apresentavam vários efeitos colaterais indesejados dos altos níveis de hormônio de crescimento. Mais notadamente, as fêmeas de porcos transgênicos eram estéreis. Além disso, animais transgênicos de ambos os sexos eram letárgicos com músculos fracos e eram mais suscetíveis a artrite e úlceras. Embora cientistas tenham a esperança de encontrar meios para superar tais efeitos colaterais, os resultados iniciais sugerem que as tentativas para melhorar as taxas de crescimento e acentuar a qualidade da carne em animais domésticos aumentando-se os níveis de hormônio de crescimento podem não ser muito eficazes.

Outros animais transgênicos estão sendo testados quanto a resistência a infecções virais. O vírus da leucose aviária (ALV) é um importante patógeno viral de galinhas, causando perdas na indústria de avicultura de US\$50 milhões a US\$100 milhões por ano. Obviamente, a disponibilidade de uma linhagem de galinhas resistente ao ALV seria de importante valor comercial. Portanto, pesquisadores produziram galinhas transgênicas portadoras de um genoma defeituoso de ALV. Essas galinhas produzem RNA viral e a proteína do envoltório viral, mas não prole de vírus. Mais importante, elas são resistentes a infecção por ALV. A síntese de

grandes quantidades da proteína do envoltório retroviral de algum modo bloqueia o ciclo reprodutivo de ALV patogênicos intactos. Essa resistência ao ALV foi transmitida a várias gerações de prole, indicando que a característica é estável. Tais resultados encorajadores sugerem que a introdução de genomas virais defeituosos em animais domésticos pode ser uma ferramenta útil na produção de genótipos resistentes a vírus.

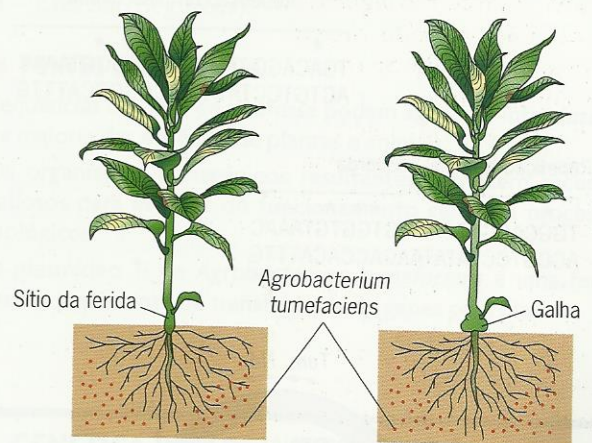
Outro emprego potencialmente importante de animais transgênicos consiste na produção e na secreção de proteínas valiosas no leite. Muitas proteínas humanas nativas contêm carboidratos ou grupos laterais de lipídeos que são adicionados pós-traducionalmente. Bactérias não contêm as enzimas que catalisam a adição destas frações a proteínas nascentes. Em tais casos, bactérias recombinantes não podem ser usadas para sintetizar o produto final. Elas irão sintetizar apenas o polipeptídeo em sua forma não modificada. Por este motivo, alguns pesquisadores começaram a explorar métodos alternativos para produzir proteínas humanas valiosas, especialmente glicoproteínas e lipoproteínas. Células de camundongo e hamster crescidas em meio de cultura hoje são comumente usadas para a produção de proteínas humanas com aplicações medicinais.

Plantas Transgênicas: O Plasmídeo Ti de *Agrobacterium tumefaciens*

Plantas têm sido geneticamente manipuladas por agricultores há décadas. Hoje em dia, entretanto, eles podem modificar diretamente o DNA de plantas e podem rapidamente adicionar genes de outras espécies a genomas de plantas por técnicas de DNA recombinante. Plantas transgênicas podem ser produzidas por vários procedimentos diferentes. Um procedimento amplamente usado, chamado de **bombardamento de microprojéteis**, envolve disparar partículas de tungstênio ou ouro revestidas com DNA em células vegetais. Outro procedimento, chamado de **eletroporação**, usa um curto choque elétrico para introduzir o DNA na célula. Entretanto, o método mais usado para gerar plantas transgênicas, pelo menos em dicotiledôneas, é provavelmente a **transformação mediada por *Agrobacterium tumefaciens***. *A. tumefaciens* é uma bactéria do solo que desenvolveu um sistema natural de engenharia genética. Ela contém um segmento de DNA que é transferido da bactéria para células vegetais.

Uma característica importante de células vegetais é sua **totipotência**, isto é, a capacidade de uma única célula de produzir todas as células diferenciadas da planta madura. Muitas células vegetais diferenciadas são capazes de se diferenciar de volta ao estado embrionário e subsequentemente se diferenciar em novos tipos de células. Assim, não há separação de células da linhagem germinativa das células somáticas como em animais superiores. Esta totipotência de células vegetais é uma grande vantagem para a engenharia genética porque permite a regeneração de plantas inteiras a partir de células somáticas individuais modificadas.

A. tumefaciens é o agente causador da doença galha de plantas dicotiledôneas. O nome refere-se a tumores que em geral se formam na coroa (junção entre a raiz e o caule) de plantas infectadas (Fig. 17.16a). Como a coroa da planta geralmente está situada na superfície do solo, é lá que a planta é mais provavelmente ferida (por exemplo, por uma abrasão do solo como em caso de um vento forte) e infectada por uma bactéria do solo tal como *A. tumefaciens*. Entretanto, *A. tumefaciens* pode infectar uma planta e induzir um tumor em qualquer local ferido (Fig. 17.16b). Após a infecção de um local ferido por *A. tumefaciens*, ocorrem dois eventos principais: (1) as células da planta começam a proliferar e formar tumores e (2) elas começam a sintetizar um derivado de arginina chamado opina. A opina sintetizada geralmente é nopalina ou octopina, dependendo da linhagem de *A. tumefaciens*. Estas opinas são catabolizadas e usadas como fontes de energia por bactérias infectantes. Linhagens de *A. tumefaciens* que induzem a síntese de nopalina podem crescer em nopalina, mas não em octopina, e vice-versa. Claramente, uma interrelação interessante desenvolveu-se entre linhagens de *A. tumefaciens* e suas plantas hospedeiras. *A. tumefaciens* é capaz de desviar os recursos metabólicos da planta hospedeira para a síntese de opinas, que não têm benefício aparente para a planta mas que dão sustentação à bactéria.



(a) Formação de uma galha na superfície do solo.



(b) Formação de tumores, ou galhas, em locais das feridas em uma folha.

Fig. 17.16 ■ A formação de tumores, ou galhas, de plantas em locais infectados por *Agrobacterium tumefaciens*.

A capacidade de *A. tumefaciens* de induzir galha em plantas é controlada por informação genética levada em um grande (cerca de 200.000 pares de nucleotídeos) plasmídeo chamado **plasmídeo Ti** por sua capacidade de induzir tumores. Dois componentes do plasmídeo Ti, o **T-DNA** e a **região vir**, são essenciais para a transformação de células vegetais. Durante o processo de transformação, o T-DNA (de DNA transferido) é retirado do plasmídeo Ti, transferido para uma célula vegetal e integrado (inserido covalentemente) no DNA da célula vegetal. Os dados disponíveis indicam que a integração do T-DNA ocorre em sítios cromossômicos aleatórios. Além disso, em alguns casos, vários eventos de integração de T-DNA ocorrem na mesma célula. Em plasmídeos Ti tipo nopalina, o T-DNA é um segmento com 23.000 pares de nucleotídeos que possui 13 genes conhecidos. Em plasmídeos Ti tipo octopina, existem dois segmentos separados de T-DNA. Para resumir, consideraremos apenas os plasmídeos Ti tipo nopalina na discussão subsequente.

A estrutura de um plasmídeo Ti típico de nopalina é mostrada na Fig. 17.17. Alguns dos genes do segmento de T-DNA do plasmídeo Ti codificam enzimas que catalisam a síntese de fitormônios (a auxina ácido indolacético e a citocinina isopentenil-adenosina). Estes fitormônios são res-

ponsáveis pelo crescimento tumoral de células em galhas. A região de T-DNA é delimitada por repetições imperfeitas de 25 pares de nucleotídeos, uma das quais deve estar presente em *cis* para excisão e transferência do T-DNA. A deleção da sequência delimitante correta bloqueia totalmente a transferência de T-DNA para as células vegetais.

A região *vir* (de virulência) do plasmídeo Ti contém os genes necessários para o processo de transferência de T-DNA. Estes genes codificam as enzimas processadoras de DNA necessárias para excisão, transferência e integração do segmento de T-DNA durante o processo de transformação. Os genes *vir* podem suprir as funções necessárias para transferência do T-DNA quando situados em *cis* ou *trans* em relação ao T-DNA. Eles são expressos em níveis muito baixos em células de *A. tumefaciens* que crescem no solo. Entretanto, a exposição da bactéria a células vegetais feridas ou exudados de células vegetais induz níveis aumentados de expressão dos genes *vir*. Este processo de indução é muito lento na bactéria, levando de 10 a 15 horas para atingir níveis máximos de expressão. Compostos fenólicos tais como acetosiringona atuam como indutores dos genes *vir*, e as taxas de transformação em geral podem ser aumentadas pela adição destes indutores a células vegetais inoculadas com *Agrobacterium*. A transformação de células vegetais pelo plasmídeo Ti de *A. tumefaciens* ocorre como ilustrado na Fig. 17.18.

Uma vez estabelecido que a região do T-DNA do plasmídeo Ti de *A. tumefaciens* é transferida para células vegetais e se torna integrada a cromossomos da planta, o uso potencial de *Agrobacterium* na engenharia genética de plantas foi óbvio. Genes exógenos podem ser inseridos no T-DNA e então transferidos para a planta com o resto do T-DNA. Este procedimento funciona muito bem. O problema é que as células vegetais transformadas que levam o T-DNA tipo selvagem perdem seu controle normal da divisão celular e formam tumores. Esta característica do T-DNA torna os plasmídeos Ti tipo selvagem incompatíveis com as metas da maioria dos experimentos de transferência gênica. Felizmente, este problema foi resolvido com a identificação dos genes no T-DNA que são responsáveis pela formação de tumor (Fig. 17.17). A deleção de um ou mais destes genes produz um plasmídeo Ti desarmado. Infelizmente, a deleção dos genes causadores de tumor também dificulta a identificação de células vegetais que receberam o T-DNA desarmado. Assim, é necessário um modo de identificar células vegetais transformadas com plasmídeos Ti desarmados – idealmente, um bom gene marcador selecionável situado dentro da região de T-DNA do plasmídeo Ti.

Um bom gene marcador selecionável é aquele que irá produzir resistência a uma droga, um antibiótico ou outro agente que cesse o crescimento de células vegetais normais. O agente seletivo deve inibir o crescimento de células vegetais ou matá-las lentamente. Agentes que rapidamente destroem células resultam na liberação de compostos fenólicos e de outras substâncias que são tóxicos para o crescimento das células restantes, de outro modo resistentes. O antibió-

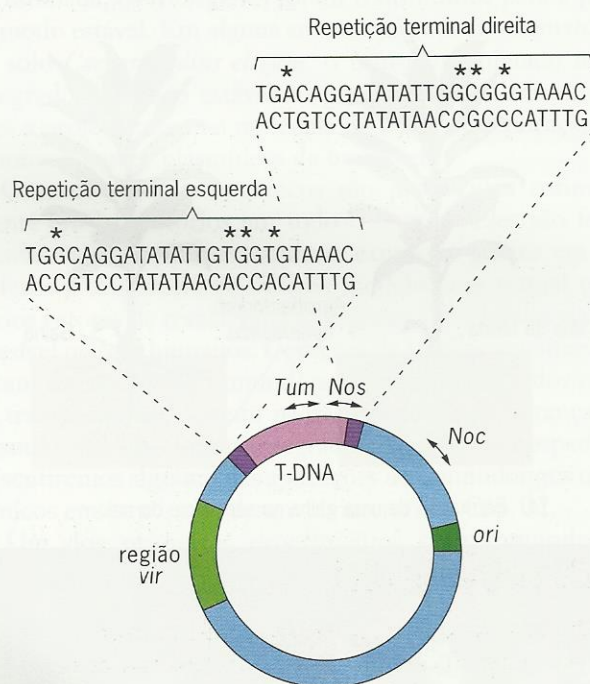


Fig. 17.17 ■ Estrutura do plasmídeo Ti de nopalina pTiC58, mostrando componentes selecionados. O plasmídeo Ti tem 210 kb de tamanho. Os símbolos usados são: *ori*, origem de replicação; *Tum*, genes responsáveis pela formação de tumores; *Nos*, genes envolvidos na biossíntese de nopalina; *Noc*, genes envolvidos no catabolismo da nopalina; *vir*, genes de virulência necessários para a transferência de T-DNA. As sequências de pares de nucleotídeos nas repetições terminais à esquerda e à direita são mostradas em cima; os asteriscos marcam os quatro pares de bases que diferem nas sequências das duas bordas.

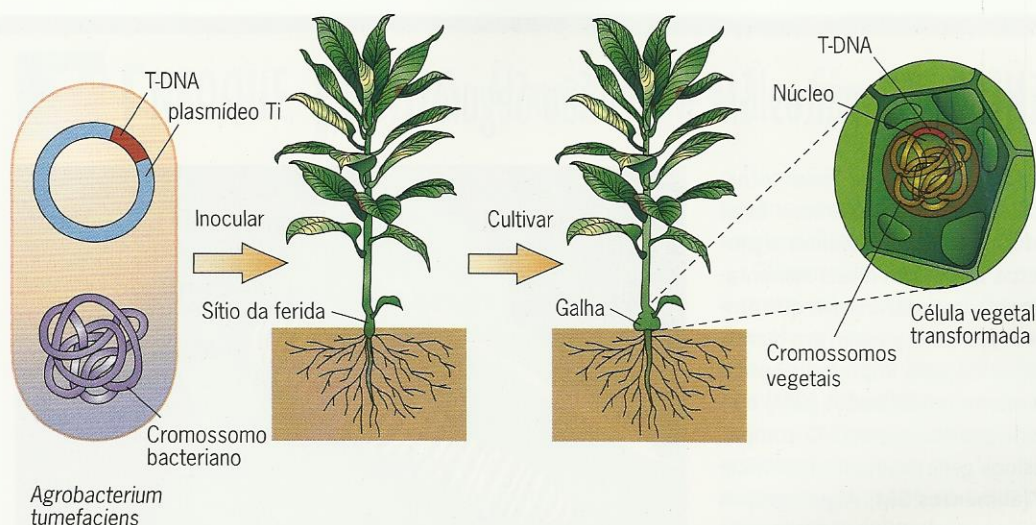


Fig. 17.18 ■ Transformação de células vegetais por *Agrobacterium tumefaciens* abrigo um plasmídeo Ti tipo selvagem. As células vegetais no tumor contêm o segmento de T-DNA do plasmídeo Ti integrado ao DNA cromossômico.

tico canamicina tem sido o agente seletivo mais amplamente usado em plantas.

O gene *kan^r* do transposon Tn5 de *E. coli* tem sido amplamente usado como um marcador selecionável em plantas. Ele codifica uma enzima chamada de neomicina-fosfotransferase tipo II (NPTII). NPTII é uma das várias enzimas procarióticas que anulam a família canamicina de antibióticos aminoglicosídicos fosforilando-os. Como as seqüências promotoras e os sinais de término de transcrição são diferentes em bactérias e plantas, o gene nativo *kan^r* de Tn5 não pode ser usado em plantas. Deve ser usada a seqüência codificante de NPTII com um promotor de planta (a 5' da seqüência codificante) e sinais de término e poliadenilação de plantas (a 3' da seqüência codificante). Tais construções com seqüências codificantes procarióticas flanqueadas por seqüências reguladoras eucarióticas são chamadas de **genes marcadores selecionáveis quiméricos**.

Seqüências reguladoras de vários genes diferentes de plantas foram usadas para construir genes marcadores quiméricos. Um gene marcador selecionável quimérico amplamente usado contém o promotor do vírus do mosaico da couve-flor (CaMV) 35S (tamanho do transcrito), a seqüência codificante NPTII e a seqüência de término da sintase de Ti nopalina (*nos*). Este gene quimérico geralmente é simbolizado como 35S/NPTII/*nos*. Os vetores Ti usados para transferir genes para plantas têm genes indutores de tumor do plasmídeo substituídos por um gene marcador selecionável quimérico tal como 35S/NPTII/*nos*. Um grande número de sofisticados vetores plasmidiais Ti de transferência de genes hoje são rotineiramente usados na transferência de genes para plantas.

As novas e poderosas ferramentas que permitem a criadores de plantas e animais produzir plantas e animais transgênicos com relativa facilidade têm uma vasta gama de aplicações. No Cap. 1, discutimos a produção de milho resistente à broca. Os transgenes mais amplamente usados são os que produzem resistência a herbicidas em cultivos agrônômicos.

Com o desenvolvimento destas e outras plantas e animais geneticamente modificados, surgiram dúvidas sobre sua segurança. De fato, uma controvérsia intensa e altamente emocional acerca da segurança de cultivos geneticamente modificados (GM) e outros alimentos está ocorrendo em muitas partes do mundo (veja Enfoque Técnico: Alimentos GM – Eles São Seguros?).

PONTOS IMPORTANTES

- Seqüências de DNA de interesse podem agora ser introduzidas na maioria das espécies de plantas e animais.
- Os organismos transgênicos resultantes fornecem recursos valiosos para estudos do funcionamento gênico e processos biológicos.
- O plasmídeo Ti de *Agrobacterium tumefaciens* é uma ferramenta importante de transferência de genes para plantas.

GENÉTICA REVERSA: DISSECANDO PROCESSOS BIOLÓGICOS POR INIBIÇÃO DA EXPRESSÃO GÊNICA

Os enfoques da genética reversa fazem uso de seqüências conhecidas de nucleotídeos para criar procedimentos de inibição da expressão de genes específicos.

A explosão de novas informações em biologia durante o século vinte ocorreu em parte devido à aplicação dos enfoques genéticos para a dissecção de processos biológicos (veja Cap. 13). O enfoque genético clássico foi identificar organismos com fenótipos anormais e caracterizar os genes mutantes responsáveis por tais fenótipos. Estudos moleculares comparativos foram então feitos em organismos mutantes e



ENFOQUE TÉCNICO: Alimentos GM – Eles São Seguros?

Tecnologias de DNA recombinante e transferência de genes tornaram rotineira a produção de plantas e animais transgênicos. Estas tecnologias foram usadas para produzir cultivos resistentes a produtos herbicidas e resistentes a insetos, arroz com níveis aumentados de vitaminas, plantas que produzem plásticos, animais que produzem proteínas humanas valiosas em seu leite, e assim em diante. Tais organismos transgênicos são referidos pela imprensa popular como **cultivos ou animais geneticamente modificados (GM)** e se tornaram o foco de uma intensa controvérsia mundial. O componente mais debatido enfoca organismos geneticamente modificados que entram na cadeia alimentar (**alimentos GM**). Alguns grupos sociais e ambientais alimentam esta controvérsia com rótulos tais como “alimento Frankenstein”, “poluição genética” e “cultivos mutantes”. Ativistas do Greenpeace tentaram bloquear o transporte internacional de cultivos geneticamente modificados, divulgando suas atividades em enormes faixas com a inscrição “Floodgate: Genetic Pollution”. Sejam ou não válidas as suas preocupações, eles criaram uma oposição importante ao desenvolvimento e ao uso de alimentos GM, especialmente na Europa, e as sociedades devem abordar tais preocupações de modo cuidadoso e razoável.

Alimentos GM são seguros? As opiniões variam de um extremo – todas as pesquisas de cultivos GM devem ser interrompidas e todos os alimentos GM devem ser destruídos – até o outro – os alimentos GM são totalmente seguros e têm um potencial fenomenal de melhorar a nutrição humana e a saúde em todo o mundo. A maioria das pessoas têm opiniões sobre alimentos GM que estão provavelmente em algum ponto entre estes extremos. Enquanto proponentes de alimentos GM enfatizam suas numerosas aplicações potenciais, oponentes questionam que alguns alimentos GM podem conter substâncias tóxicas e alergênicas e que os cultivos de GM resultarão em diminuição da biodiversidade. Alguns dos potenciais riscos e benefícios de cultivos GM são mostrados no Quadro 1.

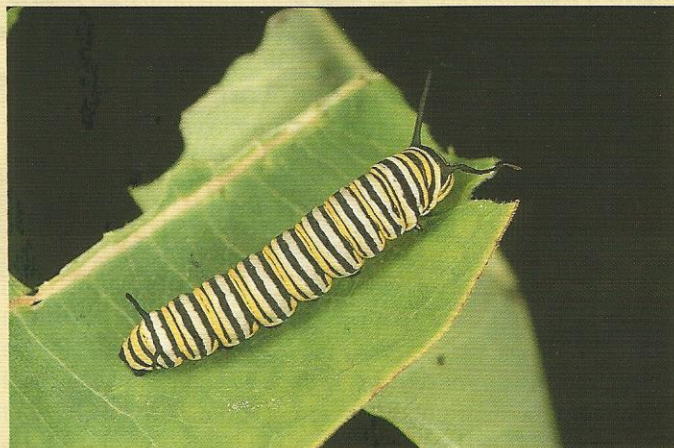


Fig. 1 ■ Uma larva da borboleta monarca alimentando-se de uma folha de asclepiadácea.

Em 1999, uma publicação na *Nature* intitulada “Transgenic Pollen Harms Monarch Larvae” de John Losey e colaboradores levantou importantes preocupações sobre os possíveis efeitos ecológicos de cultivos GM. Eles alimentaram larvas da borboleta monarca (Fig. 1) com folhas revestidas pelo pólen de plantas de milho contendo um gene que codifica uma proteína inseticida (toxina *bt*) da bactéria *Bacillus thuringiensis* (veja Fig. 1.14). Muitas das larvas morreram, resultando em preocupação sobre os efeitos dos 26 milhões de acres de cultivo de milho *bt* nos EUA naquela época. Estudos subsequentes demonstraram que esta preocupação provavelmente não tinha fundamento por vários motivos: (1) o pólen de milho não adere bem às folhas; (2) a quantidade de pólen de milho presente nas folhas próximas aos campos de milho *bt* é muito mais baixa do que a quantidade usada no estudo de Losey; (3) as linhagens mais frequentemente cultivadas de milho

Quadro 1 Riscos e Benefícios Potenciais de Cultivos GM

Riscos Potenciais	Benefícios Potenciais
<ol style="list-style-type: none"> 1. Geração de produtos tóxicos ou alergênicos. 2. Efeitos deletérios inesperados de produtos “não naturais”. 3. “Escape” de transgenes por transferência para espécies correlatas, com efeitos deletérios. 4. Polinização de cultivos “orgânicos”, reduzindo seu valor. 5. Efeitos prejudiciais na produção de mel por abelhas. 6. Efeitos prejudiciais de produtos transgênicos em outras espécies, tais como borboletas monarca. 7. Redução da biodiversidade. 8. Perdas econômicas devidas a boicotes internacionais de produtos GM. 9. Aumento da dependência por parte de fazendeiros de grandes empresas agroquímicas. 	<ol style="list-style-type: none"> 1. Alimentos com qualidade nutricional melhorada, tais como vitaminas e minerais adicionados. 2. Produção de fármacos e vacinas. 3. Cultivos resistentes a insetos, herbicidas e doenças. 4. Retardo no amadurecimento de frutos, ampliando sua duração. 5. Produção de plásticos biodegradáveis em cultivos. 6. Controle biológico de pragas, resultando em menor uso de herbicidas químicos e pesticidas. 7. Síntese de produtos que são matérias-primas valiosas para a produção de papel, detergentes, lubrificantes e similares. 8. Controle genético do tempo de floração, conteúdo das sementes e outras características importantes de cultivos.



ENFOQUE TÉCNICO: Alimentos GM – Eles São Seguros? (continuação)

bt têm níveis muito baixos da toxina em seu pólen; e (4) no campo, a maioria do pólen de milho é liberada antes que as monarchas ponham seus ovos. Subseqüentemente, Losey demonstrou que borboletas monarcha evitam folhas revestidas com pólen de milho *bt* quando botam os ovos. Além disso, o método alternativo comumente usado para controlar pestes de insetos – espalhar inseticidas químicos – é quase certamente mais prejudicial à monarcha do que ao milho *bt*. Entretanto, a preocupação do público foi muito influenciada pela publicação original de Losey, tendo sido dada menos publicidade a estudos subseqüentes.

O aspecto positivo da saga das borboletas monarcha é que aumentou a preocupação quanto aos possíveis efeitos ecológicos dos cultivos GM, o que por sua vez estimulou discussões importantes e pesquisas. Talvez a preocupação mais válida sobre cultivos GM é se eles irão diminuir a biodiversidade. O amplo crescimento de cultivos resistentes a herbicidas e o controle das pragas com um herbicida específico eliminarão algumas das plantas nativas no ecossistema? Ele será prejudicial para a vida selvagem, especialmente as aves? Vários estudos estão sendo planejados para abordar estas questões. Entretanto, mesmo tais estudos, que foram criados para obter mais informações acerca dos efeitos de

cultivos GM sobre a biodiversidade, estão sendo criticados como potencialmente daninhos pelo Greenpeace e outros grupos de ambientalistas.

Claramente, o debate emocional relativo a alimentos GM é tanto sobre valores culturais e sociais quanto sobre a ciência que ele envolve. As preocupações sobre alimentos GM devem lidar com a consideração destes fatores. Na Nova Zelândia, uma comissão de cultivos de GM tentou obter e avaliar opiniões de todos os setores da sociedade (para detalhes, veja <http://www.gmcommission.govt.nz>). O trabalho desta comissão fornece um excelente modelo para o que precisa ser feito em escala mundial. Documentos submetidos à comissão foram amplamente opostos a cultivos de GM (65% fortemente opostos). Entretanto, uma ampla pesquisa subseqüente de opinião pública observou que apenas 2% dos pesquisados citaram alimentos GM como uma de suas preocupações. Esses resultados sugerem que os opositores de alimentos GM são muito mais atuantes em expressar sua oposição, enquanto o cidadão médio está menos preocupado com os riscos potenciais associados a alimentos GM. Entretanto, duas afirmativas são certas: o debate vai continuar, e o resultado final não pode ser previsto.

tipo selvagem para determinar os efeitos das mutações. Tais estudos identificaram genes codificantes de produtos que estavam envolvidos nos processos biológicos investigados. Em alguns casos, os resultados desses estudos permitiram que biólogos determinassem a seqüência precisa de eventos ou vias pelos quais ocorre um processo. A via completa da morfogênese no bacteriófago T4 (veja Fig. 13.11) proporcionou documentação inicial do poder do enfoque da dissecação mutacional.

Durante as últimas duas décadas, as seqüências de nucleotídeos de genes e genomas inteiros tornaram-se disponíveis. Hoje em dia, geralmente conhecemos a seqüência de nucleotídeos de um gene antes de conhecer sua função. Esse conhecimento resultou em novos enfoques da dissecação genética de processos biológicos, enfoques coletivamente chamados de **genética reversa**. Os enfoques da genética reversa usam as seqüências de nucleotídeos de genes objetivando criar procedimentos para isolar mutações neutras neles ou para desligar sua expressão. A função de um gene específico pode ser deduzida estudando-se organismos sem nenhum produto funcional do gene. Nas seções seguintes deste capítulo, examinaremos três enfoques genéticos reversos diferentes: RNA anti-sentido, inserções de T-DNA e transposons e RNA de interferência. Outro enfoque de genética reversa importante, a produção de mutações “nocautes” em camundongos, é coberta no Cap. 22 como parte da discussão sobre a análise genética do desenvolvimento de vertebrados.

RNA Anti-sentido

O primeiro procedimento de genética reversa para desligar a expressão de genes específicos foi o uso de **RNA anti-sentido**. O método de RNA anti-sentido envolve a síntese de moléculas de RNA que são complementares às moléculas de mRNA produzidas por transcrição de um determinado gene. Diz-se que o mRNA normal de um gene é dotado de sentido porque leva os códons que são lidos durante a tradução para produzir a seqüência especificada de aminoácidos no produto gênico polipeptídico. Normalmente, o complemento do mRNA com sentido não conterá uma seqüência de códons que pode ser traduzida para gerar uma proteína funcional. Assim, este filamento complementar é chamado de RNA anti-sentido. Além disso, um RNA anti-sentido geralmente não contém as seqüências reguladoras necessárias para tradução. Quando moléculas de RNA anti-sentido estão presentes na mesma célula em associação a moléculas com sentido (mRNA) de um gene, as moléculas de RNA anti-sentido e de mRNA irão helicoidizar-se para formar moléculas bifilamentares de RNA, que serão degradadas ou reprimidas traducionalmente.

O enfoque de RNA anti-sentido demonstrou-se útil para dissecar vias de expressão gênica em plantas, animais e microrganismos. Ele também tem sido um sucesso comercial. O primeiro produto de uma planta geneticamente modificada aprovada para consumo humano, o tomate FlavrSavr™, foi introduzido em supermercados da Califórnia e Illinois

em 1994. O tomate FlavrSavr™, que permanece firme por mais tempo durante o processo de amadurecimento, foi produzido usando-se RNA anti-sentido para diminuir a taxa de expressão de um gene endógeno que codifica uma enzima chamada poligalacturonase para 10% do nível normal. A poligalacturonase é uma enzima que degrada as paredes celulares e amolece os tomates à medida que ficam maduros.

A maioria de nós sabe por experiência própria que tomates amadurecidos no pé são muito mais saborosos do que os que são colhidos verdes e ficam maduros a caminho do supermercado. Entretanto, tomates que amadurecem no tomateiro são muito macios para sobreviver ao manuseio necessário durante o transporte; eles são amassados facilmente. O tomate FlavrSavr™ resolveu parcialmente este problema. Eles permanecem firmes durante a maturação, permitindo que fiquem mais tempo na plantação antes de ser colhidos e enviados ao mercado.

O modo mais simples de produzir o RNA anti-sentido de um gene dentro de uma célula ou um organismo é (1) clonar o gene de interesse, (2) separar a sequência codificante do gene de seu promotor cortando o DNA com uma enzima de restrição apropriada, (3) ligar a sequência codificante a seu promotor na orientação inversa (de uma ponta a outra) e (4) introduzir a sequência codificante invertida/promotor ou **gene anti-sentido** na célula hospedeira ou organismo por transformação. O efeito de um gene que especifica o RNA anti-sentido é ilustrado na Fig. 17.19. A transcrição do gene anti-sentido irá produzir RNA transcrito anti-sentido. Esses RNA anti-sentido irão hibridizar-se com moléculas de mRNA (RNA com sentido) nas células e mediar sua degradação ou bloquear sua tradução pela via de interferência de RNA (Cap. 21). Como resultado, os RNA anti-sentido irão bloquear a expressão de genes especificando mRNA complementar. Embora os RNA anti-sentido bloqueiem a

expressão de muitos genes, eles têm pouco efeito sobre a expressão de outros genes. Assim, o uso de RNA anti-sentido para análises genéticas reversas foi amplamente substituído por enfoques usando T-DNA e mutagênese insercional de transposon ou RNA de interferência.

T-DNA e Inserções de Transposons

Em uma seção anterior deste capítulo, discutimos como o segmento de T-DNA do plasmídeo Ti de *Agrobacterium tumefaciens* é transferido para células vegetais e inserido nos cromossomos da planta (veja Fig. 17.18). Quando o T-DNA é inserido em um gene, ele perturba o funcionamento do gene. Transposons são elementos genéticos de transposição que têm a capacidade de mover-se de um local no genoma para outro local (Cap. 18). Do mesmo modo que o T-DNA de um plasmídeo Ti, um transposon irá perturbar o funcionamento de um gene no qual ele se insere (veja Fig. 13.24). Assim, T-DNA e transposons são ferramentas poderosas para análise genética reversa. Em ambos os casos, o elemento genético é usado para fazer **mutagênese insercional** – a indução de mutações nulas (em geral, chamadas de mutações “nocautes”) pela inserção de DNA exógenos nos genes. A mutagênese insercional é basicamente a mesma seja ela feita com o plasmídeo Ti ou um transposon. Assim, ilustraremos o uso da mutagênese insercional para a genética reversa discutindo o emprego de inserções de T-DNA para dissecar a função gênica na planta *Arabidopsis thaliana*.

Quando o T-DNA é transferido de *A. tumefaciens* para células vegetais, ele se integra a essencialmente todos os componentes do genoma, isto é, os T-DNA são encontrados dispersos ao longo de cada um dos cinco cromossomos de *Arabidopsis*. Portanto, se for examinada uma população

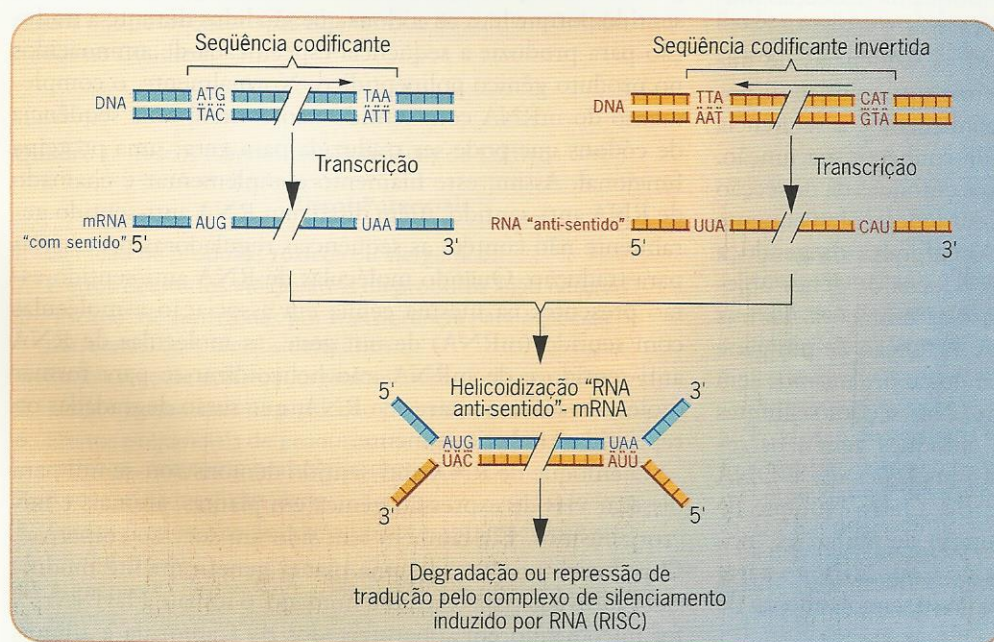


Fig. 17.19 ■ O procedimento de RNA anti-sentido para bloquear ou reduzir o nível de expressão de genes específicos. O RNA bifilamentar produzido quando o RNA anti-sentido hibridiza com o mRNA (RNA com sentido) é marcado para degradação ou repressão traducional pelo complexo de silenciamento induzido por RNA (RISC) como discutido no Cap. 21.

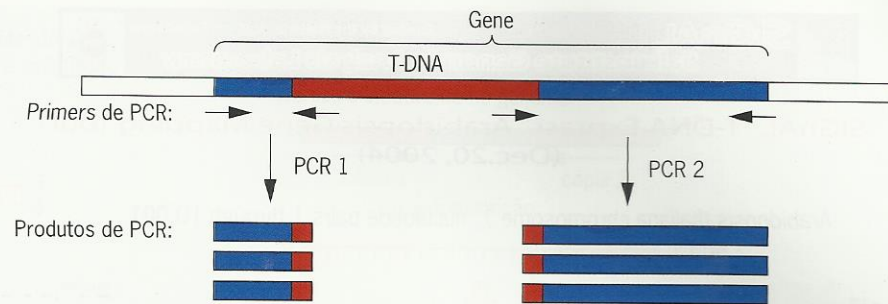


Fig. 17.20 ■ Identificação de inserções de T-DNA dentro de genes por PCR. Os *primers* de PCR que contêm as seqüências de nucleotídeos 5' → 3' nas pontas do T-DNA (chamadas seqüências delimitantes) e nas pontas do gene de interesse são usadas em PCR para amplificar os segmentos de DNA entre os insertos de T-DNA e as pontas dos genes. Como o T-DNA pode ser inserido em ambas as orientações (a mostrada e a invertida), em geral são feitas quatro amplificações de PCR – as duas mostradas e duas com os *primers* delimitantes de T-DNA da direita e da esquerda trocados. A localização exata da inserção de T-DNA é determinada por seqüenciamento de fragmentos de DNA amplificados por PCR.

grande o suficiente de plantas *Arabidopsis* transformadas, deve ser possível identificar inserções de T-DNA em todos os aproximadamente 26.000 genes desta espécie.

Durante os primeiros anos de mutagenese insercional de T-DNA em *Arabidopsis*, amostras de DNA foram isoladas de *pools* de 100 a 1.000 plantas transformadas e triadas quanto a inserções em genes específicos pela reação em cadeia da polimerase (PCR; veja Fig. 15.11). Se o DNA de um *pool* de plantas transformadas contiver uma inserção em um gene de interesse, então o DNA preparado de *pools* menores ou plantas individuais é retriado para identificar a planta contendo a inserção desejada.

O procedimento de triagem envolve a síntese de *primers* de PCR correspondentes às pontas (“seqüências delimitantes”) do T-DNA e do gene e o uso delas em pares – um *primer* de T-DNA e um *primer* de gene orientado inversamente – para amplificar os segmentos intercalares de DNA (Fig. 17.20). Na ausência de um inserto de T-DNA, nenhum fragmento de DNA será amplificado. O sítio exato do inserto de T-DNA é então determinado seqüenciando-se os produtos amplificados de PCR. O mesmo procedimento é usado para triar inserções de transposons em genes, exceto que as seqüências delimitantes de transposons são usadas como *primers* de PCR.

Hoje em dia, a identificação de linhagens de plantas transformadas com inserções de T-DNA em genes de interesse é muito mais fácil. A The National Science Foundation nos EUA e em agências governamentais em outros países tem projetos subvencionados para produzir grandes populações de plantas *Arabidopsis* transformadas por T-DNA. As seqüências de nucleotídeos dos DNA flanqueadores das inserções de T-DNA foram determinadas e usadas para identificar os sítios das inserções. Como a seqüência do genoma de *Arabidopsis* é conhecida (Cap. 16), as seqüências de DNA flanqueadoras podem ser usadas para mapear com precisão cada inserção. Em janeiro de 2005, aproximadamente 336.000 inserções de transposons e T-DNA tinham sido mapeadas no genoma de *Arabidopsis*, e o mapa dessas inserções ficou

disponível para todos colocando-as na *web*. Pesquisadores do Salk Institute em La Jolla, Califórnia, integraram o mapa das inserções de T-DNA de seu laboratório aos mapas de inserções de transposons e T-DNA caracterizados por outros grupos de pesquisadores. Seu mapa baseado em seqüência dessas inserções está disponível em <http://signal.salk.edu/cgi-bin/tdnaexpress>. Uma versão abreviada de seu mapa da ponta do cromossomo 1 é mostrada na Fig. 17.21.

Sementes das plantas transgênicas das linhagens Salk e muitas de outras coleções estão disponíveis, por solicitação, no *Arabidopsis* Biological Resource Center (ABRC) na Ohio State University. Além disso, sementes de T-DNA e linhagens de inserção de transposons caracterizadas no Versailles Genomic Resource Center (VGRC) na França, no Nottingham Arabidopsis Stock Center (NASC) na Alemanha e no Riken BioResource Center no Japão também estão disponíveis para a comunidade de pesquisas de *Arabidopsis*. Sementes das linhagens desejadas podem ser pedidas *online* no *web site* do ABRC <http://www.arabidopsis.org/abrc>. Além disso, *links* para outros *sites* com informações sobre o genoma de *Arabidopsis* estão disponíveis em <http://www.arabidopsis.org>, o *web site* do The Arabidopsis Information Resource (TAIR). Portanto, se alguém estiver interessado na função de um determinado gene de *Arabidopsis*, pode pesquisar no *web site* Salk as inserções de transposons e T-DNA neste gene. Uma vez identificadas as inserções, as sementes com as mutações de inserção desejadas podem ser solicitadas *online*. Estas grandes coleções de mutações insercionais demonstraram ser recursos valiosos para estudos da função gênica nesta planta modelo.

Interferência do RNA

Embora seus efeitos tenham sido primeiro observados em petúnias há alguns anos, a descoberta do terceiro enfoque da genética reversa, a **interferência do RNA (RNAi)**, geralmente é creditada ao trabalho de Andrew Fire, Craig Mello e colaboradores, publicado em 1998. Quando eles injetaram RNA



SIGnAL "T-DNA Express" Arabidopsis Gene Mapping Tool (Dec.20, 2004)

Arabidopsis thaliana chromosome 1, nucleotide pairs 1 through 10,001.

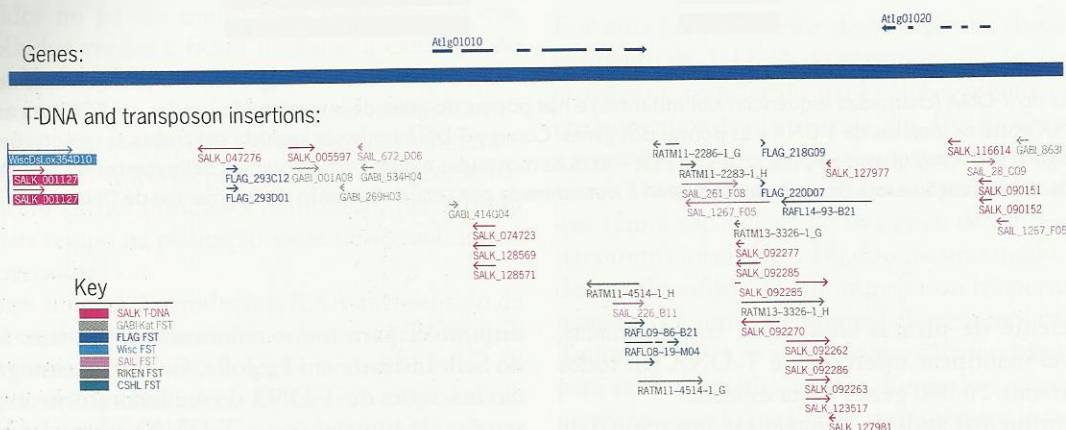


Fig. 17.21 ■ Mapa de inserções de T-DNA e transposon na região de 10 kb na ponta do cromossomo 1 em *Arabidopsis*. As posições das marcações de seqüências flanqueadoras (FSTs) são mostradas como setas abaixo do cromossomo (boxe azul-escuro). Os dados mostrados são do web site de SIGnAL (Salk Institute Genomic Analysis Laboratory), <http://signal.salk.edu/cgi-bin/tdnaexpress>. Os dois genes (*Atlg01010* e *Atlg01020*) nesta região do cromossomo 1 têm funções desconhecidas. O T-DNA e as linhagens de inserção de transposons são do Salk Institute (Salk T-DNA), da coleção alemã (GABI-Kat), da coleção francesa (FLAG), da University of Wisconsin (Wisc), da Syngenta *Arabidopsis* Insertion Library (SAIL), do Riken BioResource Center in Japan e da coleção do Cold Spring Harbor Laboratory (CSHL).

bifilamentar (dsRNA) em *C. elegans*, ele “exerceu influência sobre” (ou desligou) a expressão de genes contendo a mesma seqüência de nucleotídeos (veja Um Marco na Genética: A Descoberta do RNA de Interferência, no Cap. 21). Durante os últimos anos, a RNAi passou a ser a fronteira na biologia molecular. Hoje sabemos que o RNA bifilamentar (dsRNA) tem papéis importantes na prevenção de infecções virais, no combate da expansão de populações de elementos genéticos de transposição e na regulação da expressão gênica (veja Cap. 21). A RNAi não só é a fronteira da biologia molecular, mas tem um potencial enorme na luta contra doenças humanas. Neste capítulo, entretanto, discutiremos o uso do RNAi como ferramenta para a genética reversa, uma ferramenta com a qual estudar a função gênica e dissecar processos biológicos.

A RNAi é usada amplamente para silenciar genes – diminuir ou desligar sua expressão – em *C. elegans*, *D. melanogaster* e muitas plantas. Ela tem empregos potenciais em todas as espécies, incluindo seres humanos. A RNAi pode ocorrer de vários modos diferentes. A característica comum em todos os experimentos de RNAi é a presença de dsRNA levando pelo menos uma parte da seqüência de nucleotídeos do gene que se quer silenciar no organismo ou em células investigadas. Dois enfoques diferentes são em geral usados para atingir esta meta. Em um enfoque, o dsRNA é sintetizado *in vitro* e microinjetado no organismo (Fig. 17.22a). No

segundo enfoque, é construído um cassete de expressão gênica que leva duas cópias de pelo menos uma porção do gene de interesse em orientações inversas, que é introduzido no organismo por transformação ou transfecção (Fig. 17.22b). Quando o transgene introduzido é transcrito, ele produz uma molécula de RNA que é autocomplementar e forma uma haste e alça parcialmente bifilamentar, ou “grampo”. Em ambos os casos, o dsRNA estimula o silenciamento do gene induzido pelo RNA. O dsRNA é ligado pelo complexo de silenciamento induzido pelo RNA (RISC) e é degradado ou traducionalmente reprimido, dependendo do organismo e das células envolvidas (veja Cap. 21 para detalhes).

A RNAi é fácil de ser realizada em *C. elegans*. Estes pequenos vermes podem ser microinjetados com dsRNA, inseridos em um meio contendo o dsRNA ou alimentados com bactérias que produzem o dsRNA de interesse. Todos os três procedimentos promovem um efetivo silenciamento em *C. elegans*.

A seqüência de 99% do genoma de *C. elegans* foi publicada em dezembro de 1998. Em dois anos, grupos colaborativos de pesquisa no Reino Unido, na Alemanha, na Suíça e no Canadá usaram RNAi para sistematicamente silenciar mais de 90% dos 2.769 genes previstos no cromossomo I e mais de 96% dos 2.300 genes previstos no cromossomo III de *C. elegans*. Tais estudos deram novas informações sobre as funções de mais de 400 genes. Claramente, RNAi é uma

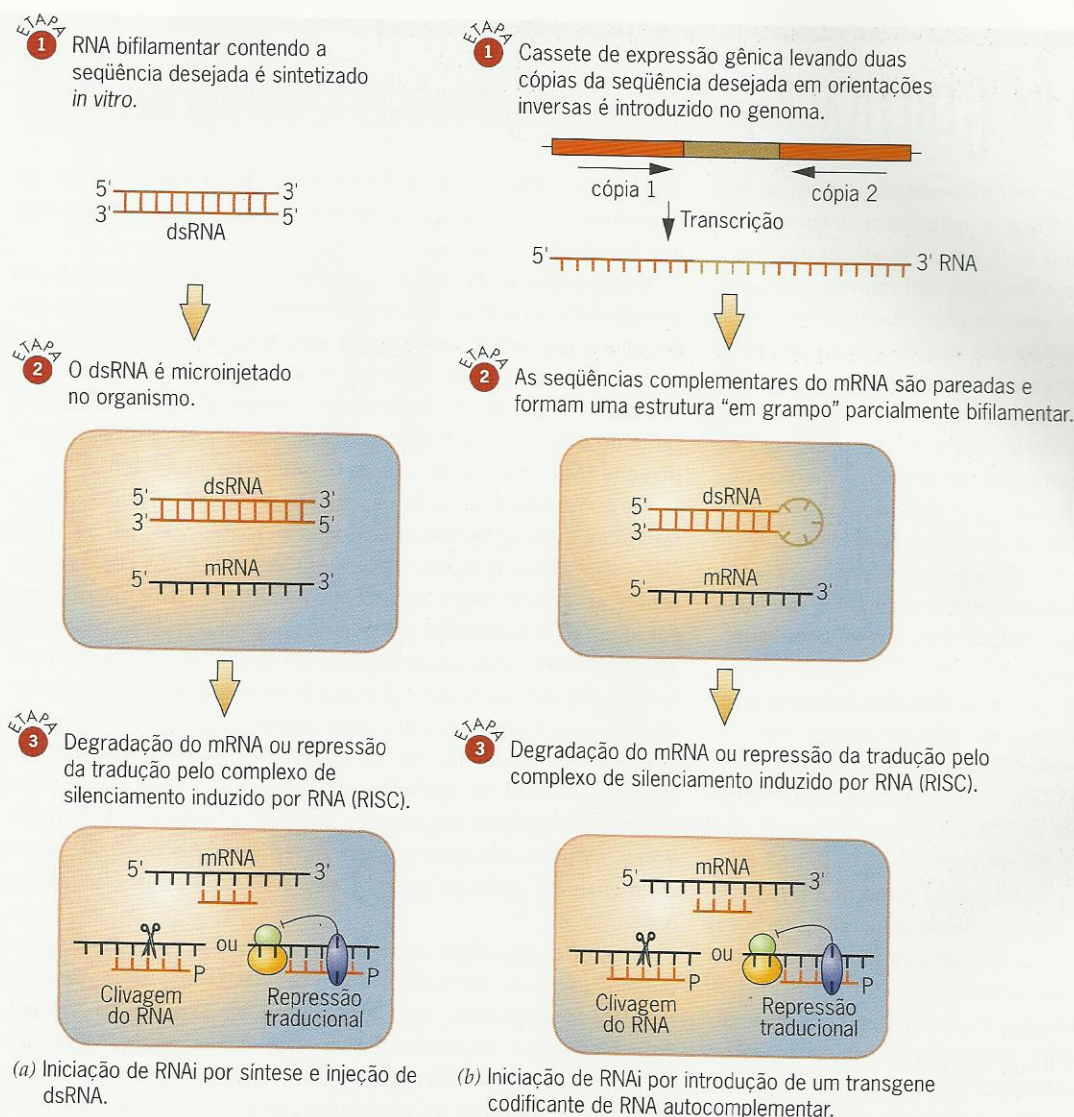


Fig. 17.22 ■ Dois procedimentos para iniciar RNAi com RNA bifilar (dsRNA). (a) Uma molécula de dsRNA contendo uma parte da sequência de nucleotídeos do gene a ser silenciado é sintetizada *in vitro* e injetada no organismo. (b) Um cassete de expressão gênica contendo duas cópias de um segmento do gene em orientações inversas é construído e introduzido no organismo pesquisado. O transcrito de RNA autocomentar forma um grampo de RNA parcialmente bifilar. Em ambos os casos, o dsRNA inicia o silenciamento do gene alvo por meio da via do complexo de silenciamento induzido por RNA (RISC), que resulta na degradação do mRNA alvo ou na repressão de sua tradução (veja Cap. 21 para detalhes).

ferramenta poderosa para dissecar processos biológicos. A RNAi usa vias naturais envolvidas na regulação da expressão gênica. Existem centenas de genes em genomas de plantas e animais que codificam **microRNA**, os quais formam dsRNA *in vivo*. No momento, conhecemos as funções regulatórias de apenas alguns desses microRNA (veja Cap. 21). Entretanto, as funções do restante dos microRNA são objeto de continuadas investigações.

Pode a RNAi ser usada para inibir a reprodução de vírus tais como o da imunodeficiência humana (HIV) ou para diminuir a expressão de oncogenes (genes causadores de câncer)? Não temos a resposta para a pergunta. Entretanto, sabemos que o mundo dos negócios está entusiasmado com as potenciais aplicações terapêuticas da RNAi. Não só as grandes empresas farmacêuticas estão investindo pesado na tecnologia da RNAi, mas muitas empresas foram formadas

especificamente para explorar a RNAi visando metas comerciais. Se as tecnologias de RNAi irão atender ou não as expectativas, só esperando para ver.

PONTOS IMPORTANTES

- Os enfoques de genética reversa usam sequências conhecidas de nucleotídeos para criar procedimentos que isolem mutações nulas de gene ou inibam a expressão gênica.
- O RNA anti-sentido, que é complementar ao RNA com sentido (mRNA), pode ser usado para desligar ou reduzir a expressão de genes individuais.
- As inserções de transposons ou T-DNA fornecem uma fonte de mutações nulas de genes.
- A interferência do RNA – bloqueio da expressão gênica com RNA bifilar – pode ser usada para dissecar processos biológicos mediante inibição das funções de genes específicos.



UM MARCO NA GENÉTICA: Repetições de Trinucleotídeos e Doença Humana

Desde a época de Archibald Garrod, milhares de distúrbios humanos hereditários foram descritos e estudados. Informações do Projeto do Genoma Humano hoje estão sendo usadas para elucidar a base molecular dessas doenças. Em muitos casos, um único gene sofreu mutação por uma simples mudança de par de bases que substitui um aminoácido por outro no produto polipeptídico do gene. Em outros casos, a mutação subjacente tem um efeito mais drástico – ela pode terminar prematuramente a síntese de polipeptídeos ou pode alterar a matriz de leitura de modo que seja produzido um polipeptídeo errado. À medida que este trabalho de detetive molecular tem progredido, geneticistas aprenderam que várias dúzias de doenças hereditárias surgem de instabilidades na composição de seqüências de genes. Mais freqüentemente, essas instabilidades envolvem regiões nas quais uma unidade de três nucleotídeos adjacentes, um trinucleotídeo, é repetida em tandem muitas vezes.

A síndrome do X frágil foi o primeiro distúrbio humano a ser associado a uma repetição instável de trinucleotídeos. Pessoas afetadas por este distúrbio exibem significativo comprometimento mental. Elas também podem apresentar anomalias faciais e comportamentais. A síndrome do X frágil ocorre em cerca de 1 em cada 4.000 homens e em cerca de 1 em cada 7.000 mulheres. Depois da síndrome Down, é a causa mais freqüente de retardo mental herdado em humanos. Estudos de heredogramas indicam que a síndrome do X frágil é causada por uma mutação dominante ligada ao X que é totalmente penetrante. Cerca de 20% dos homens hemizigotos e cerca de 30% das mulheres heterozigotas não manifestam sintomas. Entretanto, descendentes de uma pessoa assintomática podem ter os sintomas do X frágil. De fato, seu risco de ser afetado pode ser consideravelmente aumentado. Este fenômeno, primeiramente descrito por Stephanie Sherman, foi chamado de “paradoxo de Sherman”.¹

Os primeiros estudos demonstraram que a síndrome do X frágil está associada a uma anomalia citológica que é detectável em células que foram cultivadas na ausência de timidina e ácido fólico. Esta anomalia, uma constricção perto da ponta do braço longo do cromossomo X, dá a impressão de que a ponta está quase se destacando do resto do cromossomo (Fig. 1a), donde o nome cromossomo do X frágil. Análise molecular mostrou subsequente que este cromossomo contém uma repetição instável de trinucleotídeo, (CGG)_n, no sítio frágil.² Essa repetição está situada na região 5'-não traduzida de um gene designado como *FMR-1*, de gene 1 de retardo mental por X frágil (Fig. 1b). O produto protéico

deste gene, referido como FMRP, acumula-se nos dendritos de neurônios, que são longas extensões do corpo celular neuronal que fazem conexões com outras células.

FMRP é uma proteína de ligação ao RNA. Ela é encontrada em complexos com mRNA e outros componentes do aparelho de tradução e pode ter um papel no transporte de moléculas de mRNA ou na regulação de sua tradução, possivelmente interagindo com microRNA (veja Cap. 21). Pessoas com a síndrome do X frágil aparentemente não produzem esta proteína. Assim, a falta de expressão do gene *FMR-1* parece ser a causa da síndrome do X frágil.

De que modo a perda de expressão de FMRP está relacionada à repetição instável de trinucleotídeos na região 5' do gene *FMR-1*? Genes *FMR-1* normais, isto é expressos, contêm 6 a 59 cópias desta repetição. Em contraste, genes *FMR-1* anormais, isto é, não expressos, que são encontrados em pessoas portadoras da síndrome do X frágil, contêm mais de 200 cópias. De algum modo, um aumento no número de repetições de trinucleotídeos interfere na expressão do gene *FMR-1*. Uma hipótese atualmente em investigação é que o aumento do número de repetições provoca modificação química do DNA no promotor do gene *FMR-1*. Esse promotor é altamente metilado em pessoas que têm a síndrome do X frágil. Vários estudos mostraram que a hipermetilação do DNA, especialmente ao redor de promotores, silencia a expressão gênica (veja Cap. 21).

E quanto aos casos em que o número de repetições de trinucleotídeo está entre 60 e 200? Pessoas com este número de repetições geralmente não apresentam a síndrome do X frágil. Entretanto, seus descendentes sim. Diz-se que genes *FMR-1* com 60 a 200 cópias da repetição de trinucleotídeo estão em um estado de pré-mutação. O aumento do número de cópias da repetição pode criar o estado de mutante, ou mutação completa. Tais aumentos podem ocorrer na linhagem germinativa feminina, mas aparentemente não na linhagem germinativa masculina. Durante replicação do DNA, a DNA-polimerase pode “deslizar” ou “tropeçar” quando passa por uma região contendo várias repetições em tandem (Fig. 2). Após sistemas de reparo corrigirem as estruturas em grampo resultantes, a região repetida pode ser aumentada significativamente. Este tipo de mecanismo pode, portanto, explicar por que a região repetida é instável de geração para geração. Entretanto, não sabemos por que essa instabilidade deve ocorrer só em mulheres. Em qualquer caso, a expansão das repetições de trinucleotídeos na região 5' do gene *FMR-1* explica o paradoxo de Sherman. Pessoas que levam um alelo de pré-mutação do gene não manifestam sintomas de X frágil. Entretanto, se a região repetida neste alelo se amplia o suficiente durante a replicação do DNA, pode ser criado um alelo mutante. Qualquer prole que herde este alelo mutante exibirá a síndrome do X frágil.

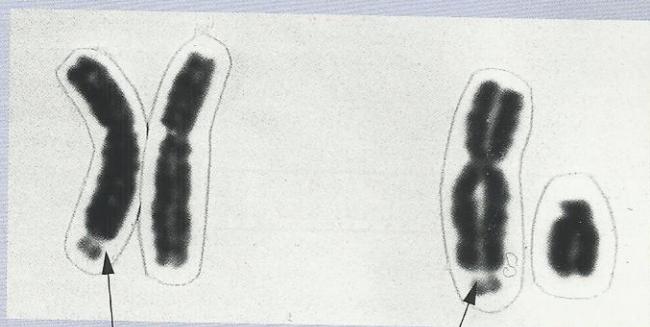
Em um ano após a descoberta da repetição instável de trinucleotídeos no gene *FMR-1*, outro distúrbio neurodegenerativo, a atrofia muscular espinobulbar (também conhecida como

¹Sherman, S. L., N. E. Morton, P. A. Jacobs e G. Turner. 1984. The marker (X) syndrome: a cytogenetic and genetic analysis. *Annals of Human Genetics* 48: 21-37.

²Fu, Y-H, D. P. A. Kuhl, A. Pizzuti, M. Pieretti, J. S. Sutcliffe, S. Richards, A. J. M. H. Verkerk, J. J. A. Holden, R. G. Fenwick, Jr., S. T. Warren, B. A. Oostra, D. L. Nelson e C. T. Caskey. 1991. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 67: 1047-1058.



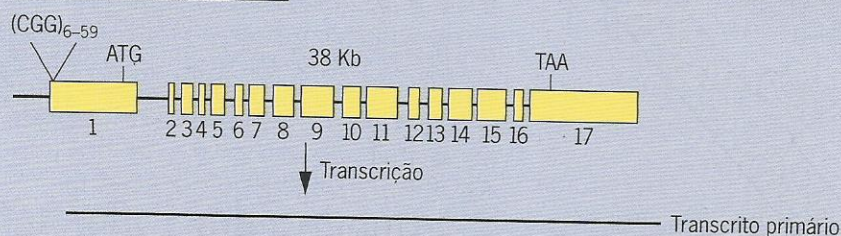
UM MARCO NA GENÉTICA (continuação)



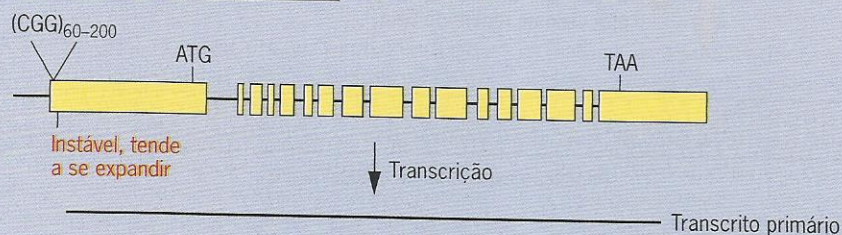
Sítio frágil

(a)

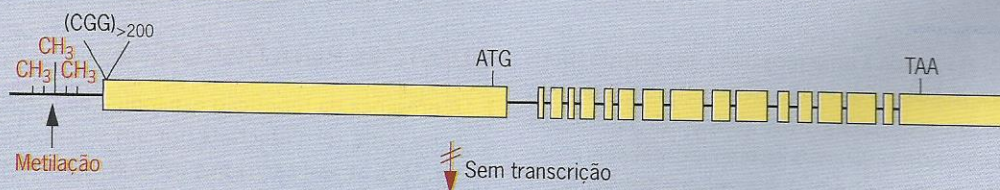
Alelos normais de *FMR-1*



Pré-mutação de alelos *FMR-1*



Alelos *FMR-1* mutantes



(b)

Fig. 1 ■ (a) O cromossomo X frágil e um X normal de uma mulher (à esquerda) e o cromossomo X frágil e um Y normal de um homem (à direita). (b) A localização e o número de repetições de trinucleotídeos CGG em alelos normal (em cima), pré-mutação (no centro), e mutante (embaixo) do gene *FMR-1*. Os promotores dos alelos mutantes são muito metilados, o que bloqueia a transcrição.



UM MARCO NA GENÉTICA (continuação)

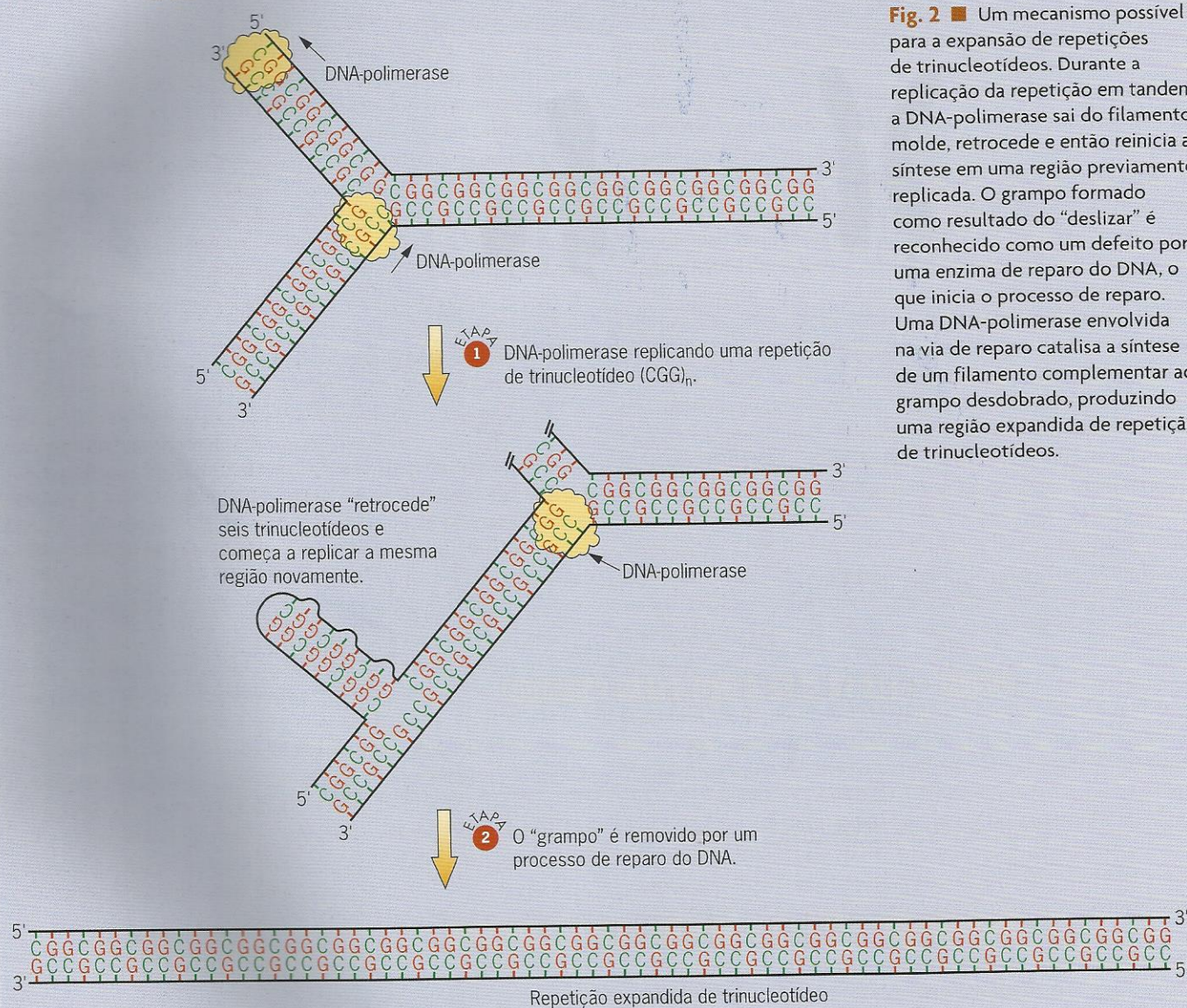


Fig. 2 ■ Um mecanismo possível para a expansão de repetições de trinucleotídeos. Durante a replicação da repetição em tandem, a DNA-polimerase sai do filamento molde, retrocede e então reinicia a síntese em uma região previamente replicada. O grampo formado como resultado do "deslizar" é reconhecido como um defeito por uma enzima de reparo do DNA, o que inicia o processo de reparo. Uma DNA-polimerase envolvida na via de reparo catalisa a síntese de um filamento complementar ao grampo desdobrado, produzindo uma região expandida de repetição de trinucleotídeos.

doença Kennedy), foi associada a uma repetição instável de trinucleotídeos, desta vez (CAG)_n.³ Foi então demonstrado que muitos outros distúrbios neurodegenerativos resultam de repetições expandidas de trinucleotídeos. O mais conhecido deles é a doença Huntington. Mutações que envolvem repetições instáveis de trinucleotídeos portanto parecem ser um tipo significativo de defeito genético em nossa espécie.

³La Spada, A. R., E. M. Wilson, D. B. Lubahn, A. E. Harding, e H. Fischbeck. 1991. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 352: 77-79.

QUESTÕES PARA DISCUSSÕES

1. A presença da repetição expandida de trinucleotídeos em um gene pode ser detectada por um simples teste de DNA (veja Fig. 17.2). Se você acha que pode ser portador de um alelo pré-mutação do gene *FMR-1*, você faria o teste? Caso sim, com quem você compartilharia os resultados?
2. Qual poderia ser o significado evolutivo de repetições de trinucleotídeos? Como elas influenciariam a composição e a estrutura de genomas?

Exercícios Básicos Ilustram a análise genética básica.

1. Como foram usados os polimorfismos de comprimento de fragmentos de restrição (RFLP) na pesquisa do gene mutante que causa a doença Huntington (HD)?

Resposta: As equipes de pesquisa de HD triaram membros de duas grandes famílias quanto a ligação entre RFLP e o gene HD. Elas encontraram um RFLP no cromossomo 4 que estava muito ligado ao gene HD (4% de recombinação).

2. Uma vez estabelecida uma forte ligação entre o gene HD e o RFLP no cromossomo 4, qual foi o próximo passo das equipes de pesquisa em sua procura pelo gene HD?

Resposta: Em seguida, elas prepararam um mapa de restrição detalhado desta região (englobando 500 kb) do cromossomo 4 (veja Fig. 17.1).

3. Como as equipes de pesquisa identificam genes candidatos dentro de uma região mapeada do cromossomo 4?

Resposta: Elas usaram clones de cDNA para identificar os segmentos codificantes ou éxons dos genes na região e triar bibliotecas genômicas quanto a clones superpostos de éxons. As seqüências dos cDNA e DNA genômicos foram então comparadas para deduzir as estruturas éxon-intron dos genes na região mapeada.

4. Como as equipes de pesquisa de HD determinaram qual dos genes candidatos era o gene HD?

Resposta: Elas seqüenciaram os genes candidatos de indivíduos com HD e os membros não afetados de suas famílias e

procuraram anomalias estruturais nos genes de pessoas afetadas. Seus resultados mostraram que um gene, hoje chamado de *huntingtina*, contém uma repetição de trinucleotídeo, (CAG)_n, que estava presente em 11 a 34 cópias de pessoas não afetadas e em 42 a mais de 100 cópias em pessoas não afetadas. Elas identificaram esta repetição expandida de trinucleotídeos nos alelos de *huntingtina* de membros afetados de 72 famílias diferentes, deixando poucas dúvidas de que *huntingtina* é o gene HD.

5. Qual o valor do conhecimento da seqüência de nucleotídeos do gene *huntingtina* para consultores genéticos?

Resposta: O conhecimento da seqüência de nucleotídeos do gene *huntingtina* forneceu a consultores um teste diagnóstico simples e preciso da presença de alelos mutantes do gene. *Primers* oligonucleotídicos de seqüências flanqueadoras da região de repetição de trinucleotídeos do gene podem ser usados para amplificar este segmento do gene, e pode ser determinado o número de repetições de trinucleotídeos por eletroforese em gel de poliácridamida (veja Fig. 17.2). Como resultado, pessoas em risco de transmitir o gene mutante podem ser testadas quanto a sua presença antes de começar uma família. Se o gene mutante está presente em um dos genitores, células fetais ou mesmo uma única célula de um embrião pré-implantação de oito células podem ser testados quanto a sua presença. Assim, consultores genéticos são capazes de fornecer às famílias em risco para o distúrbio informação precisa a respeito da presença do gene em pessoas que planejam suas famílias, em células fetais e mesmo em embriões de oito células.

Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. Ataxia espinocerebelar (tipo 1) é uma doença neurológica progressiva com início ocorrendo tipicamente entre os anos 30 e 50. A neurodegeneração resulta da perda seletiva de neurônios específicos. Embora não seja compreendido por que ocorre morte neuronal seletiva, sabe-se que a doença é causada pela expansão de uma repetição de trinucleotídeos CAG, com alelos normais contendo cerca de 28 cópias e alelos mutantes que possuem 43 a 81 cópias do trinucleotídeo. Considerando as seqüências de nucleotídeos em ambos os lados da região repetida, como você testaria a presença da região repetida expandida de trinucleotídeos responsável pela ataxia espinocerebelar tipo 1?

Resposta: O teste de DNA para ataxia espinocerebelar (tipo 1) seria similar ao teste para o alelo *huntingtina* descrito na Fig. 17.2. Primeiro você faria *primers* de PCR correspondentes a seqüências de DNA em ambos os lados da região CAG repetida. Esses *primers* seriam usados para amplificar a região CAG repetida desejada do DNA genômico da pessoa que está sendo testada por PCR. Então, os tamanhos das regiões de trinucleotídeos repetidas seria determinada medindo-se os tamanhos dos produtos de PCR por eletroforese em gel. Qualquer gene com menos de 30 cópias da repetição CAG seria considerado um alelo normal, enquanto a presença de um gene com 40 ou mais

cópias do trinucleotídeo seria diagnóstica dos alelos mutantes que causam ataxia espinocerebelar.

2. Suponha que você fez o teste de DNA para ataxia espinocerebelar em uma mulher com 25 anos de idade cuja mãe morreu da doença. Os resultados foram positivos para a mutação de ataxia. A mulher e seu marido desejam ter seu próprio filho biológico, mas não querem o risco de transmitir o gene defeituoso para nenhum de seus filhos. Quais são as opções deles?

Resposta: Suas opções irão depender de suas convicções religiosas e morais. Uma possibilidade envolve o uso de amniocentese ou biopsia coriônica para obter células fetais no início da gravidez, fazendo teste de DNA quanto à região com expansão de trinucleotídeos responsável por ataxia espinocerebelar em células fetais e só permitindo a continuação da gestação se o gene defeituoso não estiver presente. Outra possibilidade é o uso de fertilização *in vitro*. O teste de DNA para ataxia é então feito em uma célula de um pré-embrião de oito células, e o pré-embrião só é implantado se o teste para o gene defeituoso de ataxia for negativo. Uma terceira opção pode estar disponível no futuro, ou seja, um método efetivo para tratamento da doença antes do início da neurodegeneração, talvez por terapia de substituição do gene.

Questões e Problemas Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 17.1.** O que são ilhas de CpG? Que valor têm as ilhas de CpG na clonagem posicional de genes humanos?
- 17.2.** Por que o gene mutante que causa a doença Huntington é denominado *huntingtina*? Por que este gene deve ser renomeado no futuro?
- 17.3.** Como a sequência de nucleotídeos do gene *CF* é usada para se obter informações sobre a estrutura e função de seu produto gênico?
- 17.4.** Como a caracterização do gene *CF* e seu produto pode proporcionar o tratamento de fibrose cística por terapia gênica de células somáticas? Que obstáculos devem ser superados antes que a fibrose cística possa ser tratada de forma bem-sucedida por terapia gênica?
- 17.5.** A distrofia miotônica (MD), que ocorre em cerca de 1 em 8.000 pessoas, é a forma mais comum de distrofia muscular em adultos. A doença, que é caracterizada por degeneração muscular progressiva, é causada por um gene mutante dominante que contém uma região com a repetição CAG expandida. Os alelos tipo selvagem do gene *MD* contêm 5 a 30 cópias do trinucleotídeo. Os alelos mutantes contêm 50 a mais de 2.000 cópias da repetição CAG. A sequência completa de nucleotídeos do gene *MD* está disponível. Imagine um teste diagnóstico para o gene mutante responsável por distrofia miotônica que possa ser feito usando-se DNA genômico de neonatos, células fetais obtidas por amniocentese e células únicas de pré-embriões de oito células produzidos por fertilização *in vitro*.
- 17.6.** Em humanos, a ausência de uma enzima chamada purina-nucleosídeo-fosforilase (PNP) resulta em uma grave imunodeficiência de células T similar à da doença de imunodeficiência combinada grave (SCID). A deficiência de PNP apresenta um padrão de herança autossômica recessiva, e o gene que codifica PNP humana já foi clonado e sequenciado. Deficiência de PNP seria boa candidata a tratamento com terapia gênica? Imagine um procedimento para o tratamento de deficiência de PNP por terapia gênica de células somáticas.
- 17.7.** Proteínas humanas agora podem ser produzidas em bactérias tais como *E. coli*. Entretanto, não podemos simplesmente introduzir um gene humano em *E. coli* e esperar que ele seja expresso. Que passos devem ser dados para construir uma linhagem de *E. coli* que produza uma proteína de mamífero tal como o hormônio de crescimento humano?
- 17.8.** Você construiu um gene sintético que codifica uma enzima que degrada o herbicida glifosato. Você deseja introduzir seu gene sintético em plantas *Arabidopsis* e testar as plantas transgênicas quanto a resistência ao glifosato. Como você pode produzir uma planta transgênica *Arabidopsis* que abriga seu gene sintético por uma transformação mediada por *A. tumefaciens*?
- 17.9.** Um *locus* de VNTR humano contém uma repetição em tandem (TAA)_n, onde n pode estar entre 5 e 15. Quantos alelos deste *locus* você esperaria encontrar na população humana?
- 17.10.** Os *fingerprints* de DNA tiveram papéis importantes em muitos julgamentos recentes de estupro e assassinato. O que é um *fingerprint* de DNA? Que papéis as impressões de DNA desempenham nestes casos forenses? Em alguns casos, geneticistas estiveram envolvidos com dados de *fingerprints* de DNA que foram usados imprópriamente. Quais são algumas de suas preocupações e como tais preocupações podem ser apropriadamente abordadas?
- 17.11.** Especialistas forenses, em sua maioria, concordam em que o *fingerprint* de DNA de amostras de sangue obtidas em cenas de crimes e em objetos pessoais pode fornecer evidência convincente para condenações por assassinato. Entretanto, a defensoria às vezes questiona, com sucesso, que o manuseio impróprio de amostras de sangue resulta em contaminação das amostras. Que problemas a contaminação de amostras de sangue apresenta na interpretação de *fingerprints* de DNA? Você esperaria que tais erros resultem na condenação de uma pessoa inocente ou na absolvição de uma pessoa culpada?
- 17.12.** Um grupo de corpos foi encontrado queimado em uma floresta. A polícia suspeita de que eles podem incluir a família Jones desaparecida (os pais e duas crianças). Eles extraíram o DNA de ossos e examinaram os genes (usando PCR) *A* e *B*, que se sabe conter repetições em tandem de trincas de tamanho variável. Eles também analisaram o DNA de dois outros homens. Os resultados são mostrados adiante onde os números indicam o número de cópias de uma repetição em tandem em um alelo particular; por exemplo, o homem 1 tem um alelo com 8 e outro alelo com 9 cópias de uma repetição em tandem no gene *A*.

	Gene A	Gene B
homem 1	8/9	5/7
homem 2	6/8	5/5
homem 3	7/10	7/7
mulher	8/8	3/5
criança 1	7/8	5/7
criança 2	8/8	3/7

A mulher poderia ser a mãe de ambas as crianças? Por que sim ou por que não? Qual homem, se algum, poderia ser o pai da criança 1?

- 17.13.** Os *fingerprints* de DNA mostrados em seguida foram preparados usando-se DNA genômico de células sangüíneas obtidas de uma mulher, sua filha, e três homens que dizem ser o pai da menina. Com base nas impressões de DNA, o que pode ser determinado sobre a paternidade neste caso?



- 17.14.** O plasmídeo Ti contém uma região chamada T-DNA. Por que esta região é chamada de T-DNA e qual o seu significado?
- 17.15.** A geração de plantas transgênicas usando-se a transformação mediada por *A. tumefaciens* em geral resulta em muitos locais de inserção. Estes sítios geralmente variam no nível de expressão de transgene. Que enfoques você poderia usar para determinar se plantas transgênicas levam ou não mais que um transgene e, caso sim, se os transgenes foram inseridos em cromossomos?
- 17.16.** Vetores retrovirais desarmados podem ser usados para introduzir genes em animais superiores, incluindo seres humanos. Que vantagens os vetores retrovirais têm em relação a outros tipos de vetores de transferência de genes? Quais as desvantagens?
- 17.17.** Camundongos transgênicos hoje são rotineiramente produzidos e estudados em laboratórios de pesquisa em todo o mundo. Como são produzidos camundongos transgênicos? Que tipos de informação podem ser obtidos de estudos feitos em camundongos transgênicos? Essa informação tem alguma importância para a prática da medicina? Caso sim, qual?
- 17.18.** O que é um RNA anti-sentido? Que emprego o RNA anti-sentido tem na pesquisa genética? na agricultura? Como pode ser produzida uma célula ou um organismo que produz um RNA anti-sentido específico?
- 17.19.** Muitas proteínas humanas valiosas contêm componentes de carboidratos ou lipídeos que são adicionados pós-transcricionalmente. Bactérias não contêm as enzimas necessárias para adicionar estes componentes a produtos primários de tradução. Como tais proteínas podem ser produzidas usando-se animais transgênicos?
- 17.20.** Richard Meagher e colaboradores clonaram uma família de 10 genes que codificam actinas (um componente importante do citoesqueleto) em *Arabidopsis thaliana*. Os 10 produtos do gene de actina são similares, em geral diferindo por apenas alguns aminoácidos. Assim, as seqüências codificantes dos 10 genes também são muito similares, de modo que a região codificante de um gene faz hibridização cruzada com as regiões codificantes dos outros nove genes. Em contraste, as regiões não codificantes dos 10 genes são bem divergentes. Meagher supôs que os 10 genes de actina apresentam padrões temporais e espaciais bem diferentes de expressão. Você foi contratado por Meagher para testar esta hipótese. Crie experimentos que lhe permitam determinar o padrão temporal e espacial de expressão de cada um dos 10 genes de actina em *Arabidopsis*.
- 17.21.** O primeiro camundongo transgênico resultou de microinjeção de ovócitos fertilizados com um vetor de DNA similar ao diagramado na Fig. 17.13 exceto que ele contém um promotor para o gene de metalotioneína de mamífero ligado ao gene de *HGH*. O camundongo transgênico resultante mostrou níveis elevados de HGH em tecidos e órgãos à exceção da hipófise – por exemplo, no coração, no pulmão e no fígado; na hipófise, ocorreu atrofia. Como a produção de HGH em animais transgênicos poderia ser mais bem regulada, com expressão restrita à hipófise?
- 17.22.** Como os enfoques da genética reversa para dissecar processos biológicos difere de enfoques da genética clássica?
- 17.23.** De que modos os procedimentos de RNA anti-sentido e silenciamento gênico por RNAi são iguais? diferentes?
- 17.24.** Como os enfoques de mutagenese insercional diferem de outros enfoques da genética reversa?
- 17.25.** A mutagenese insercional é uma poderosa ferramenta tanto em plantas quanto em animais. Entretanto, quando se faz mutagenese insercional em grande escala, que vantagem importante as plantas têm em relação aos animais?
- 17.26.** Discutimos os efeitos desafortunados de mutagenese insercional nos dois garotos que desenvolveram leucemia após tratamento da doença de imunodeficiência combinada grave ligada ao X por terapia gênica (veja Fig. 17.9). Como esta consequência da terapia gênica pode ser evitada no futuro? Você acredita que o uso de terapia gênica com células somáticas para tratar doenças humanas possa ainda ser feito 100% livre de risco? Por quê? Por que não?
- 17.27.** Um filamento de um gene de *Arabidopsis thaliana* tem a seguinte seqüência de nucleotídeos:
- ```
atgagtgcacggaggaggaagaagagcgtgaacggagggtcaccggcgcaaa
caatcttgatgatcgagatctagcttccggaagtgtgaagcttccaccg
gtcgggaacgagctgttatcaagagtgccgatgaaagatgatatgcaaaag
gaagctatcgaaatcgccatctccgcgtttgagaagtacagtgtggagaag
gatatagctgagaatataagaaggaggtttgacaagaaacatggtgctacttggcattg
cattgttggtcgcaacttgggtcttctatgtaacgatgagacaaaccatttcgttact
tctacctgcaccagaaagctgtgctgctcttcaagtcgggttaa
```



A(s) função(ões) deste gene ainda é(são) incerta(s). (a) Como você pode usar um RNA anti-sentido para estudar sua(s) função(ões)? (b) Como a mutagenese insercional pode ser usada para investigar a(s) função(ões) do gene? (c) Crie um experimento usando interferência de RNA para sondar a(s) função(ões) do gene.

- 17.28. Vamos examinar o *web site* (<http://signal.salk.edu/cgi-bin/td-naexpress>) do Salk Institute Genome Analysis Laboratory

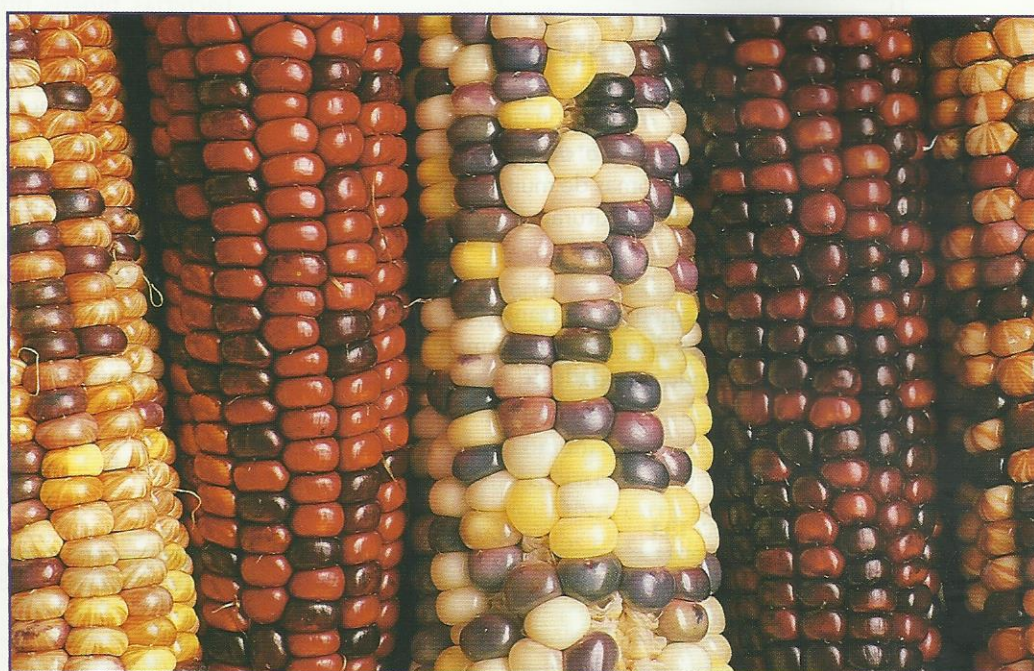
para ver se alguma de suas linhagens de T-DNA tem inserções no gene mostrado na pergunta anterior. No *web site* SIGNAL, desça até "Blast" e digite a sequência no box. O mapa resultante mostrará a localização das inserções mapeadas de T-DNA relativas à localização do gene (retângulo verde no topo). As setas azuis à direita do topo permitem-lhe enfocar a região curta contendo o gene ou regiões relativamente longas do cromossomo 4 de *Arabidopsis*. Existe alguma inserção de T-DNA no gene em questão? perto do gene?



# Elementos Genéticos de Transposição

## Tópicos

- Elementos de Transposição: Uma Visão Geral
- Elementos de Transposição em Bactérias
- Recortar e Colar Transposons em Eucariontes
- Retrovírus e Retrotransposons
- Elementos de Transposição em Humanos
- O Significado Genético e Evolutivo de Elementos de Transposição



Variação de cores entre grãos de milho. Estudos da base genética desta variação resultaram na descoberta dos elementos de transposição.

## Milho: Um Cultivo Básico com uma Herança Cultural

O plantio de milho é um dos cultivos mais importantes do mundo. O cultivo do milho começou há pelo menos 5.000 anos na América Central. Na época em que Cristóvão Colombo chegou ao Novo Mundo, o cultivo do milho já se havia expandido desde o norte (Canadá) até o sul (Argentina). Os nativos das Américas do Norte e do Sul desenvolveram muitas variedades diferentes de milho, cada uma adaptada a condições particulares. Eles desenvolveram variedades que tinham grãos coloridos – vermelhos, azuis, amarelos, brancos e púrpura – e associaram cada cor a um valor estético ou religioso especial. Para as pessoas do sudoeste norte-americano, por exemplo, o milho azul é considerado sagrado, e cada uma das quatro direções cardinais é representada por uma cor particular de milho. Alguns grupos consideram os grãos com listras e pontos como sinais de força e vigor.



Os padrões coloridos que vemos nas espigas de milho também têm um importante significado científico. Pesquisas modernas mostraram que as listras e pontos em grãos de milho são o resultado de um fenômeno genético chamado de *transposição*. No genoma do milho, como dentro dos genomas da maioria dos organismos, geneticistas encontraram seqüências de DNA que podem mover-se de uma posição para outra. Estes *elementos de transposição* – ou, mais simplesmente, *transposons* – constituem uma fração apreciável do genoma. Em milho, por exemplo, eles contribuem com mais de metade de todo o DNA. Quando elementos de transposição se movem de um local para outro, eles podem quebrar cromossomos ou mutar genes. Assim, tais elementos têm um profundo significado genético.

## ELEMENTOS DE TRANSPOSIÇÃO: UMA VISÃO GERAL

Elementos de transposição, ou transposons, são encontrados nos genomas de muitos tipos de organismos; eles são estrutural e funcionalmente diversos.

O mundo dos elementos de transposição é rico em diversidade. Muitos tipos diferentes de elementos foram identificados em uma grande variedade de organismos, incluindo

bactérias, fungos, protistas, plantas e animais. Estes elementos são componentes destacados de genomas – por exemplo, mais de 40% do genoma humano – e eles claramente exercem papéis na modelação da estrutura dos cromossomos e na modulação da expressão dos genes. Neste capítulo, exploraremos a diversidade estrutural e comportamental de tipos diferentes de elementos de transposição e investigaremos seu significado genético e evolutivo.

Embora cada tipo de elemento de transposição tenha suas próprias características especiais, a maioria deles pode ser classificada em uma dentre três características com base na maneira como eles são transpostos (Quadro 18.1).

QUADRO 18.1

### Categorização dos Elementos de Transposição pelo Mecanismo de Transposição

| Categoria                                                                                            | Exemplos                                                            | Organismo Hospedeiro |
|------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------|----------------------|
| I. Transposons de recortar e colar                                                                   |                                                                     |                      |
|                                                                                                      | Elementos IS (p. ex., IS50)                                         | Bactérias            |
|                                                                                                      | Transposons compostos (p. ex., Tn5)                                 | Bactérias            |
|                                                                                                      | Elementos <i>Ac/Ds</i>                                              | Milho                |
|                                                                                                      | Elementos <i>P</i>                                                  | <i>Drosophila</i>    |
|                                                                                                      | Elementos <i>mariner</i>                                            | <i>Drosophila</i>    |
|                                                                                                      | Elementos <i>hobo</i>                                               | <i>Drosophila</i>    |
|                                                                                                      | Elementos <i>Tc1</i>                                                | Nematódeos           |
| II. Transposons replicativos                                                                         | Elementos Tn3                                                       | Bactérias            |
| III. Retrotransposons                                                                                |                                                                     |                      |
| A. Elementos tipo retrovírus (também chamados longas repetições terminais, ou LTR, retrotransposons) |                                                                     |                      |
|                                                                                                      | Ty1                                                                 | Leveduras            |
|                                                                                                      | <i>copia</i>                                                        | <i>Drosophila</i>    |
|                                                                                                      | <i>gypsy</i>                                                        | <i>Drosophila</i>    |
| B. Retroposons                                                                                       |                                                                     |                      |
|                                                                                                      | Elementos <i>F</i> , <i>G</i> e <i>I</i>                            | <i>Drosophila</i>    |
|                                                                                                      | Retroposons específicos de telômeros ( <i>HeT-A</i> , <i>TART</i> ) | <i>Drosophila</i>    |
|                                                                                                      | LINE (p. ex., <i>L1</i> )                                           | Humanos              |
|                                                                                                      | SINE (p. ex., <i>Alu</i> )                                          | Humanos              |



Na primeira categoria, a transposição é feita excisando-se um elemento de sua posição em um cromossomo e inserindo-o em outra posição. Os eventos de excisão e inserção são catalisados por uma enzima chamada transposase, que é codificada pelo próprio elemento. Geneticistas referem-se a este mecanismo como *transposição de cortar e colar* porque o elemento é fisicamente cortado de um sítio em um cromossomo e colado em outro sítio, que pode ser até mesmo em um cromossomo diferente. Chamaremos os elementos nesta categoria de **transposons de cortar e colar**.

Na segunda categoria, a transposição é feita por um processo que envolve replicação do DNA do elemento de transposição. Uma transposase codificada pelo elemento medeia uma interação entre o elemento e um sítio potencial de inserção. Durante essa interação, o elemento é replicado e uma cópia dele é inserida em um novo local; uma cópia também permanece no local original. Como existe um ganho de uma cópia do elemento, geneticistas referem-se a este mecanismo como *transposição replicativa*. Chamaremos tais elementos de **transposons replicativos**.

Na terceira categoria, a transposição é feita por um processo que envolve a inserção de cópias de um elemento que são sintetizadas a partir do RNA do elemento. Uma enzima chamada de transcriptase reversa usa o RNA do elemento como um molde para sintetizar moléculas de DNA, que são então inseridas em novos sítios cromossômicos. Como esse mecanismo reverte o sentido usual pelo qual flui a informação genética nas células, isto é, flui do RNA para o DNA e não do DNA para o RNA, geneticistas referem-se a ele como *retrotransposição*. Nós chamaremos os elementos nesta categoria de **retrotransposons**. Alguns dos elementos que são transpostos deste modo estão relacionados aos retrovírus e são conseqüentemente chamados de *elementos similares a retrovírus*. Outros elementos que participam da retrotransposição são simplesmente chamados de *retroposons*.

Consideraremos muitos elementos de transposição diferentes neste capítulo, cada um com sua própria história peculiar. O Quadro 18.1 categoriza estes elementos de acordo com seus mecanismos de transposição. Os transposons de recortar e colar são encontrados tanto em procariontes quanto em eucariontes. Os transposons replicativos são encontrados apenas em procariontes, e os retrotransposons são encontrados apenas em eucariontes.

### ■ PONTOS IMPORTANTES

- Um transposon de cortar e colar é removido de uma posição genômica e inserido em outra por uma enzima, a transposase, que é geralmente codificada pelo próprio transposon.
- Um transposon replicativo é copiado durante o processo de transposição.
- Um retrotransposon produz moléculas de RNA que sofrem transcrição reversa em moléculas de DNA. Essas moléculas de DNA são subseqüentemente inseridas em novas posições genômicas.

## ELEMENTOS DE TRANSPOSIÇÃO EM BACTÉRIAS

Os transposons bacterianos movem-se dentro do cromossomo bacteriano e vários tipos de plasmídeos e entre eles.

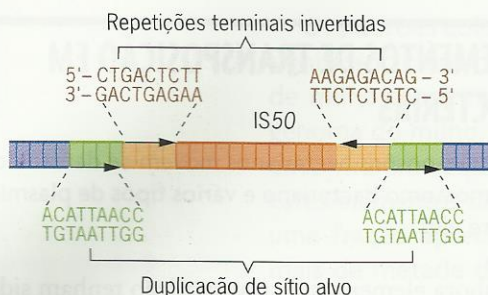
Embora elementos de transposição tenham sido originalmente descobertos em eucariontes, os transposons bacterianos foram os primeiros a serem estudados em nível molecular. Existem três tipos principais: as seqüências de inserção, ou elementos IS, os transposons compostos e os elementos Tn3. Estes três grupos de transposons diferem em tamanho e estrutura. Os elementos IS são os mais simples, contendo apenas genes que codificam proteínas envolvidas em transposição. Os transposons compostos e os elementos Tn3 são mais complexos, contendo alguns genes que codificam produtos não relacionados ao processo de transposição.

### Elementos IS

Os transposons bacterianos mais simples são as **seqüências de inserção** ou **elementos IS**, assim chamados porque podem inserir-se em muitos sítios diferentes em cromossomos bacterianos e plasmídeos. Elementos IS foram primeiro detectados em algumas mutações *lac*<sup>-</sup> de *E. coli*. Tais mutações têm a propriedade incomum de reverter ao tipo selvagem em alta taxa. Análises moleculares revelaram que essas mutações instáveis possuíam um DNA extra nos genes *lac* ou perto deles. Quando o DNA dos revertentes tipo selvagem destas mutações foram comparados com os das próprias mutações, foi descoberto que o DNA extra havia sido perdido. Assim, tais mutações geneticamente instáveis foram causadas por seqüências de DNA que se haviam inserido em genes de *E. coli*, e a reversão para o tipo selvagem foi causada pela excisão destas seqüências. Seqüências de inserção similares foram encontradas em muitas outras espécies bacterianas.

Os elementos IS são organizados de modo compacto. Tipicamente, eles consistem em menos de 2.500 pares de nucleotídeos e contêm apenas genes cujos produtos estão envolvidos em promover ou regular a transposição. Muitos tipos distintos de elementos IS já foram identificados. O menor, IS1, tem 768 pares de nucleotídeos de tamanho. Cada tipo de elemento IS é demarcado por seqüências idênticas, ou quase idênticas, curtas em suas pontas (Fig. 18.1). Como essas seqüências terminais estão sempre em orientação inversa uma em relação a outra, elas são chamadas de **repetições terminais invertidas**. Seus tamanhos variam de 9 a 40 pares de nucleotídeos. Repetições terminais invertidas são características da maioria dos tipos de transposons mas não todos. Quando nucleotídeos nestas repetições estão mutados, o transposon geralmente perde sua capacidade de mover-se. Tais mutações demonstram, portanto, que repetições terminais invertidas têm um papel importante no processo de transposição.

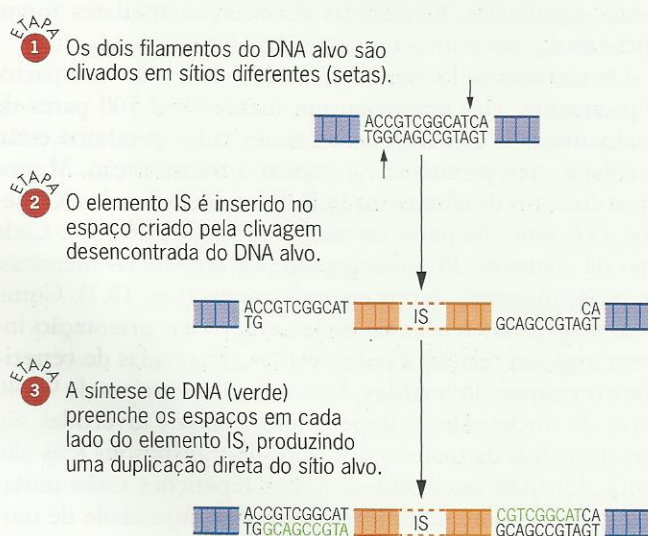




**Fig. 18.1** ■ Estrutura de um elemento de inserção IS50 mostrando suas repetições terminais invertidas e a duplicação de sítios alvos. As repetições terminais invertidas são imperfeitas porque o quarto par de nucleotídeos de cada ponta é diferente.

Pelo menos alguns elementos IS codificam uma proteína que é necessária para transposição. Esta proteína, chamada **transposase**, parece ligar-se perto das pontas do elemento, onde corta ambos os filamentos do DNA. A clivagem do DNA nestes sítios remove o elemento do cromossomo ou plasmídeo, de modo que ele pode ser inserido em uma nova posição na mesma molécula ou em uma molécula diferente de DNA. Os elementos IS são, portanto, transposons de cortar e colar. Quando elementos IS são inseridos em cromossomos ou plasmídeos, eles criam uma duplicação de parte da sequência de DNA no sítio da inserção. Uma cópia da duplicação está situada em cada lado do elemento. Estas curtas (2 a 13 pares de nucleotídeos) sequências de repetição direta, chamadas de **duplicações de sítio alvo**, parecem surgir de clivagem desencontrada da molécula bifilamentar de DNA (Fig. 18.2).

Um cromossomo bacteriano pode conter várias cópias de um determinado tipo de elemento IS. Por exemplo, 6 a 10 cópias de IS1 são encontradas no cromossomo de *E. coli*.



**Fig. 18.2** ■ A produção de duplicações de sítio alvo pela inserção de um elemento IS.

Plasmídeos também podem conter elementos IS. O plasmídeo F, por exemplo, tipicamente tem pelo menos dois elementos IS diferentes, IS2 e IS3. Quando um determinado elemento IS reside tanto em um plasmídeo quanto em um cromossomo, ele cria a oportunidade de recombinação homóloga entre moléculas diferentes de DNA. Tal recombinação parece ser responsável pela integração do plasmídeo F no cromossomo de *E. coli* (Cap. 8).

Tanto o cromossomo de *E. coli* quanto o plasmídeo F são moléculas de DNA circular. Quando o plasmídeo e o cromossomo se recombinam em uma região de homologia – por exemplo, em um elemento IS comum a ambos –, o plasmídeo menor é integrado ao cromossomo maior, criando uma única molécula circular. Tais eventos de integração produzem linhagens Hfr capazes de transferir seus cromossomos durante a conjugação. Essas linhagens variam no sítio de integração do plasmídeo F porque os elementos IS que medeiam a recombinação ocupam diferentes posições cromossômicas em linhagens diferentes de *E. coli*, um resultado de sua capacidade de se transpor. Devido a seu papel em formar linhagens Hfr, os elementos IS potencializam a troca de genes cromossômicos entre linhagens diferentes de bactérias. Essa troca cria variabilidade genética, que, juntamente com mutações, é a base da mudança evolutiva em populações de bactérias.

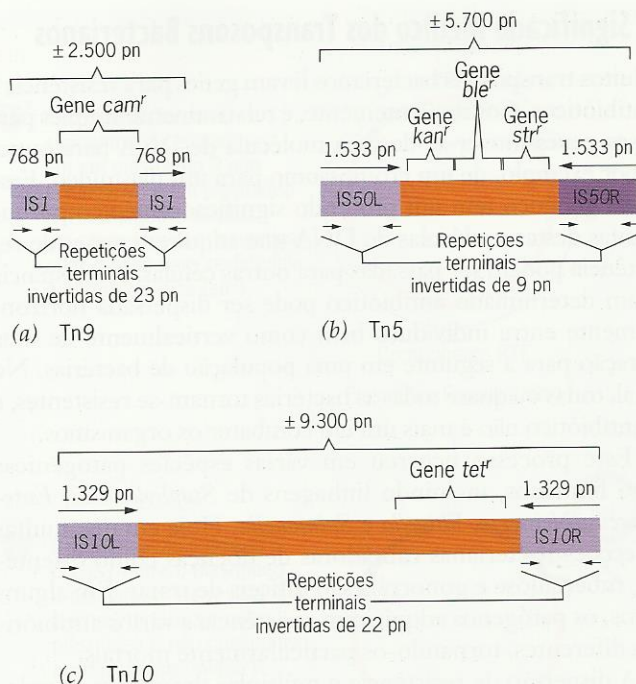
## Transposons Compostos

**Transposons compostos**, que são transposons bacterianos de cortar e colar indicados pelo símbolo Tn, são criados quando dois elementos IS são inseridos um perto do outro. A região entre eles pode então ser transposta pela ação conjunta dos elementos flanqueadores. Os dois elementos IS “capturam” uma sequência de DNA que de outro modo é imóvel, e lhe conferem a capacidade de mover-se. A Fig. 18.3 dá três exemplos. Em Tn9, os elementos flanqueadores IS estão na mesma orientação com relação um ao outro, enquanto em Tn5 e Tn10 a orientação é invertida. A região entre os elementos IS em cada um destes transposons contém genes que não guardam relação com transposição. De fato, em todos os três transposons, os genes entre os elementos flanqueadores IS conferem resistência a antibióticos.

Às vezes, os elementos flanqueadores IS em um transposon composto não são bem idênticos. Por exemplo, no Tn5, o elemento à direita, chamado de IS50R, é capaz de produzir uma transposase para estimular a transposição, mas o elemento à esquerda, chamado de IS50L, não é capaz. Esta diferença é devida a uma mudança em um único par de nucleotídeo que impede que IS50L especifique a transposase ativa.

Tn5 também ilustra outra característica dos transposons compostos: seu movimento é regulado (Fig. 18.4). Quando uma bactéria é infectada com um bacteriófago não lítico que leva Tn5 em seu cromossomo, a frequência de transposição de Tn5 é intensamente reduzida se a bactéria infectada já contiver uma cópia de Tn5. Essa redução significa que



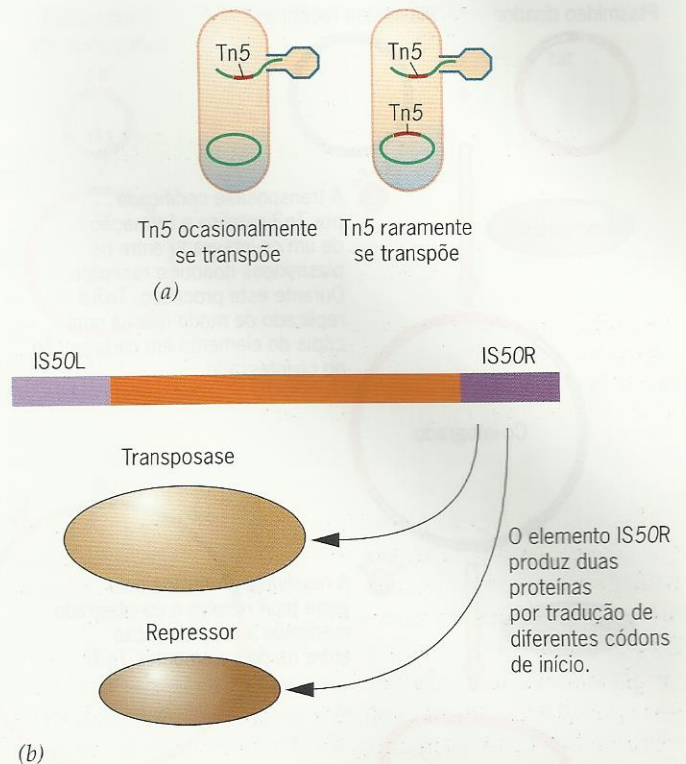


**Fig. 18.3** ■ A organização genética de transposons compostos. São indicados a orientação e o tamanho (em pares de nucleotídeos, pn) das seqüências constituintes. (a) Tn9 consiste em dois elementos IS1 flanqueadores de um gene para resistência a cloranfenicol. (b) Tn5 consiste em dois elementos IS50L e IS50R flanqueadores dos genes de resistência a canamicina, bleomicina e estreptomicina. (c) Tn10 consiste em dois elementos IS10L e IS10R flanqueadores de um gene para resistência a tetraciclina.

o transposon residente inibe a transposição de um transposon que chega, possivelmente sintetizando um repressor. Análises realizadas por Michael Syvanen, William Reznikoff e seus colegas mostraram que esta hipótese é correta. O elemento IS50R de Tn5 de fato produz duas proteínas. Uma, a transposase, catalisa transposição, enquanto a outra, uma transposase truncada criada pela tradução a partir de um códon de início dentro do gene da transposase, impede a transposição. Como a proteína mais curta é mais abundante, a transposição de Tn5 tende a ser reprimida.

### Elementos Tn3

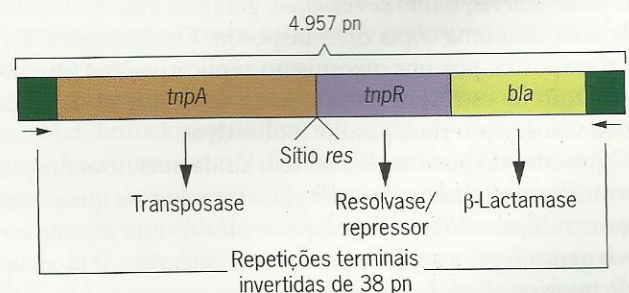
Os elementos neste grupo de transposons são maiores do que os elementos IS e geralmente contêm genes que não são necessários para transposição, uma característica dos transposons compostos. Entretanto, ao contrário dos transposons compostos, os elementos Tn3 não têm elementos IS em cada uma de suas pontas. Em vez disso, os elementos Tn3 apresentam repetições invertidas de 38 a 40 pares de nucleotídeos de comprimento em suas terminações. Os elementos Tn3 também produzem duplicações de sítio alvo quando se inserem no DNA.



**Fig. 18.4** ■ Regulação de Tn5. (a) Infecção de *E. coli* com bacteriófagos que levam Tn5. As bactérias que já possuem uma cópia de Tn5 reprimem a transposição. (b) Base genética da regulação de Tn5. Uma das proteínas produzidas por IS50R é uma transposase que catalisa a transposição, mas a outra é um repressor que inibe a transposição. O efeito do repressor geralmente prevalece.

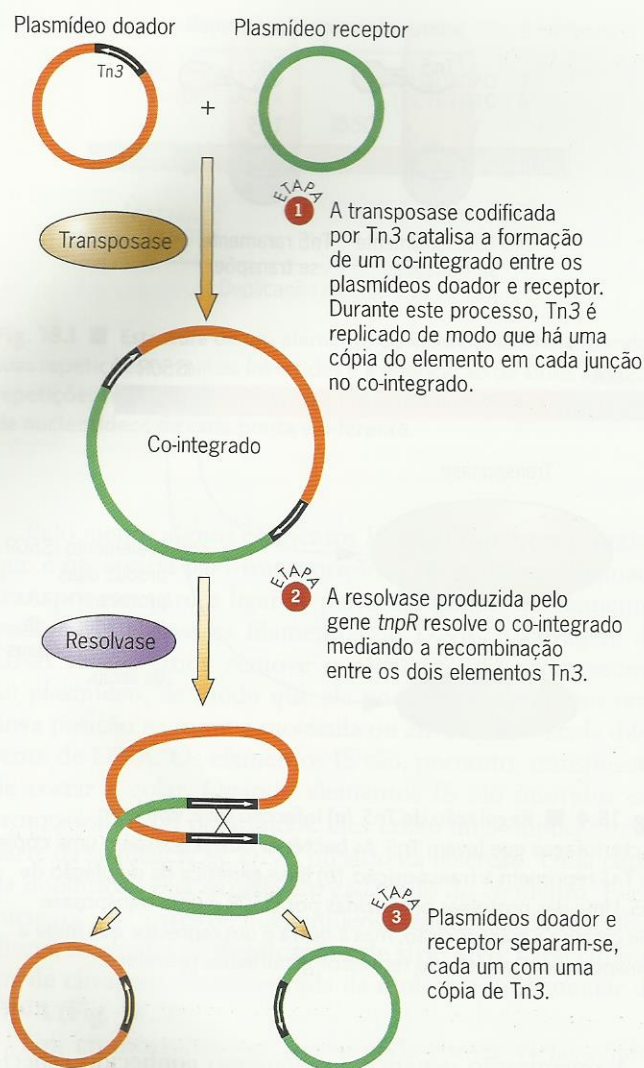
A organização genética do elemento conhecido especificamente como Tn3 é mostrada na Fig. 18.5. Existem três genes, *tnpA*, *tnpR* e *bla*, codificando, respectivamente, uma transposase, uma resolvase/repressor e uma enzima chamada beta-lactamase. A beta-lactamase confere resistência ao antibiótico ampicilina, e as outras duas proteínas têm papéis importantes na transposição.

A transposição de Tn3 ocorre em dois estágios (Fig. 18.6). Primeiro, a transposase medeia a fusão de duas moléculas circulares, uma levando Tn3 e a outra não. A estrutura resultante é chamada um **co-integrado**. Durante este processo,



**Fig. 18.5** ■ Organização genética de Tn3. Os tamanhos das seqüências de DNA são dados em pares de nucleotídeos (pn).





**Fig. 18.6** ■ Transposição de Tn3 mediante a formação de um co-integrado.

o transposon é replicado, e uma cópia é inserida em cada junção no co-integrado; os dois elementos Tn3 no co-integrado são orientados no mesmo sentido. No segundo estágio de transposição, a resolvase codificada por *tnpR* medeia um evento de recombinação sítio-específico entre os dois elementos Tn3. Este evento ocorre em uma sequência em Tn3 chamada *res*, o *sítio de resolução*, gerando duas moléculas, cada uma com uma cópia do transposon. Os elementos Tn3 são transpostos por um mecanismo replicativo que envolve a formação da estrutura co-integrada de transição. Eles são classificados como transposons replicativos.

O produto do gene *tnpR* Tn3 tem ainda uma nova função, reprimir a síntese tanto das proteínas transposase quanto resolvase. A repressão ocorre porque o sítio *res* está situado entre os genes *tnpA* e *tnpR*. Ligando-se a este sítio, a proteína *tnpR* interfere na transcrição de ambos os genes, deixando seus produtos em pouca quantidade crônica. Como resultado, o elemento Tn3 tende a permanecer imóvel.

## O Significado Médico dos Transposons Bacterianos

Muitos transposons bacterianos levam genes para resistência a antibióticos. Consequentemente, é relativamente simples para esses genes mover-se de uma molécula de DNA para outra, – por exemplo, de um cromossomo para um plasmídeo. Esse fluxo genético tem um profundo significado médico porque muitas destas moléculas de DNA que adquirem genes de resistência podem ser passadas para outras células. A resistência a um determinado antibiótico pode ser dispersada horizontalmente entre indivíduos bem como verticalmente de uma geração para a seguinte em uma população de bactérias. No final, todas ou quase todas as bactérias tornam-se resistentes, e o antibiótico não é mais útil em combater os organismos.

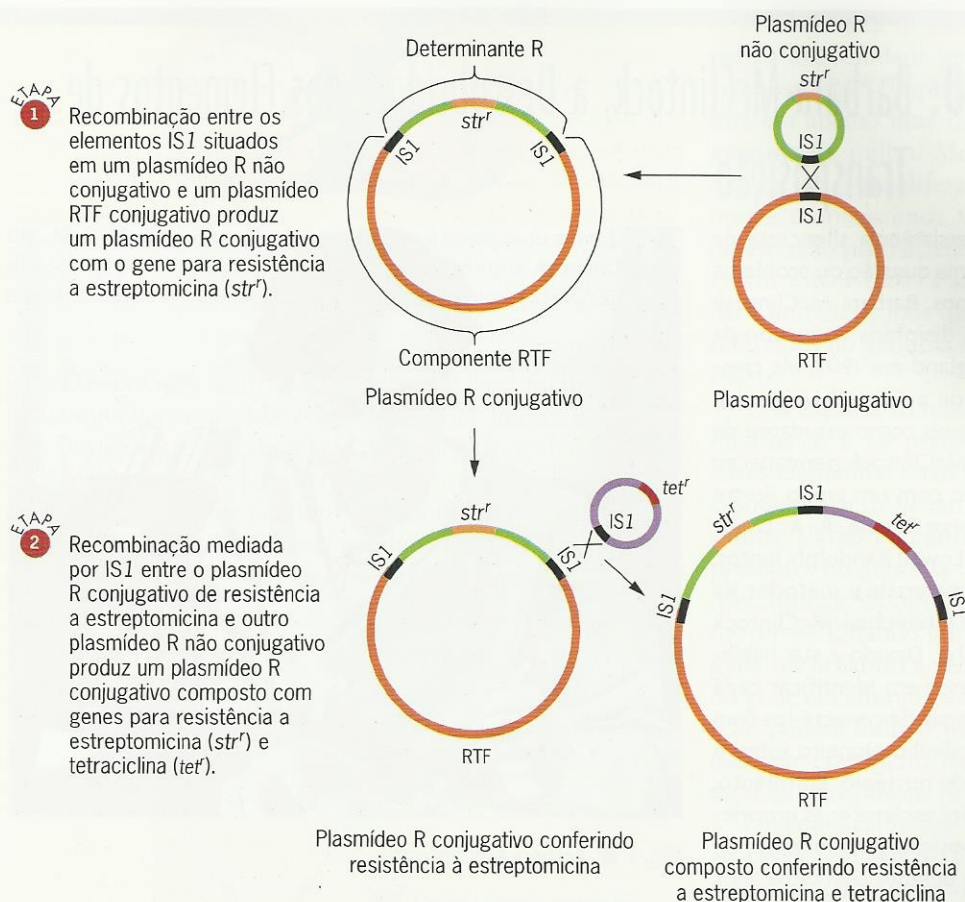
Este processo ocorreu em várias espécies patogênicas para humanos, incluindo linhagens de *Staphylococcus*, *Enterococcus*, *Neisseria*, *Shigella* e *Salmonella*. Hoje em dia, muitas infecções bacterianas causadoras de doenças como disenteria, tuberculose e gonorréia são difíceis de tratar. Em alguns casos, os patógenos adquiriram resistência a vários antibióticos diferentes, tornando-os particularmente mortais.

A dispersão de resistência a múltiplas drogas em populações bacterianas foi acelerada pela evolução de **plasmídeos R conjugativos** que levam genes de resistência. Esses plasmídeos têm dois componentes. Um, chamado de **fator de transferência de resistência**, ou **RTF**, contém os genes necessários para transferência conjugativa entre células; o outro, chamado de **determinante R**, contém os genes para resistência a antibióticos. Em geral, tais genes de resistência são levados por um transposon, ou um grupo de transposons, que foi inserido no plasmídeo (Fig. 18.7). Plasmídeos R conjugativos podem ser transferidos rapidamente entre células em uma população bacteriana. Os plasmídeos conjugativos podem, de fato, ser passados de uma espécie para outra, até mesmo entre tipos celulares bem diferentes – por exemplo, entre cocos e bacilos. Assim, uma vez que resistência a várias drogas se tenha desenvolvido em uma parte do reino microbiano, ela pode espalhar-se para outras partes com relativa facilidade.

### PONTOS IMPORTANTES

- Sequências de inserção (elementos IS) são transposons de cortar e colar que residem em cromossomos bacterianos e plasmídeos.
- Os transposons compostos consistem em dois elementos IS flancadores de uma região que contém um ou mais genes de resistência a antibióticos.
- Tn3 é um transposon replicativo que se transpõe por fusão temporária de moléculas de DNA para um co-integrado. Quando o co-integrado é resolvido, cada uma das moléculas de DNA constituintes emerge com uma cópia de Tn3.
- Os transposons bacterianos são demarcados por repetições terminais invertidas. Quando elas se inserem em uma molécula de DNA, criam uma duplicação no local de inserção.
- Plasmídeos conjugativos podem mover transposons que contêm genes para resistência a antibióticos de uma bactéria para outra.





**Fig. 18.7** ■ Evolução de plasmídeos conjugativos que levam genes para resistência a antibióticos.

## RECORTAR E COLAR TRANSPOSONS EM EUKARIOTES

Os elementos de transposição foram descobertos mediante análise de instabilidades genéticas em milho; análises genéticas também revelaram elementos de transposição em *Drosophila*.

Geneticistas encontraram muitos tipos diferentes de transposons em eucariotes. Tais elementos variam em tamanho, estrutura e comportamento. Alguns são abundantes no genoma, outros são raros. Nas seções seguintes, discutimos alguns dos transposons eucarióticos que são transpostos por um mecanismo de recortar e colar. Todos esses elementos têm repetições invertidas em suas pontas e criam duplicações em sítio alvo quando se inserem em moléculas de DNA. Alguns codificam uma transposase que catalisa o movimento do elemento de uma posição para outra.

### Elementos *Ac* e *Ds* em Milho

Os elementos *Ac* e *Ds* em milho foram descobertos pelo trabalho pioneiro de Barbara McClintock (veja Enfoque Técnico: Barbara McClintock, a Descobridora dos Elementos

de Transposição). Por meio de análise genética, McClintock mostrou que as atividades destes elementos são responsáveis pelas listras e pontos nos grãos de milho. Muitos anos depois, Nina Federoff, Joachim Messing, Peter Starlinger, Heinz Saedler, Susan Wessler e seus colaboradores isolaram os elementos e determinaram sua estrutura molecular.

McClintock descobriu os elementos *Ac* e *Ds* estudando a quebra de cromossomos. Ela usou marcadores genéticos que controlavam a cor dos grãos de milho para detectar os eventos de quebra. Quando um determinado marcador era perdido, McClintock deduzia que o segmento cromossômico no qual estava situado também tinha sido perdido, uma indicação de que havia ocorrido uma quebra. A perda de um marcador foi detectada por uma mudança na cor da aleurona, a camada mais externa do endosperma triplóide de grãos de milho.

Em um conjunto de experimentos, o marcador genético que McClintock seguia era um alelo do locus *C* no braço curto do cromossomo 9. Como esse alelo, *C<sup>d</sup>*, é um inibidor dominante da coloração de aleurona, qualquer grão que o possuía era incolor. McClintock fertilizou plantas *CC* com pólen de *C<sup>d</sup>C<sup>d</sup>*, produzindo grãos nos quais o endosperma era *C<sup>d</sup>CC*. (O endosperma triplóide recebe dois alelos do genitor feminino e um do genitor masculino.) Embora McClintock tenha descoberto que a maioria desses grãos eram incolores,





## ENFOQUE TÉCNICO: Barbara McClintock, a Descobridora dos Elementos de Transposição

Avanços científicos podem surgir da persistência silenciosa de uma única pessoa, alguém que enfoque uma questão ou problema e proceda a sua pesquisa por muitos anos. Barbara McClintock era tal pessoa (Fig. 1). Sua longa vida foi devotada ao estudo da genética do milho. Nascida em New England em 1902, ela cresceu na cidade de New York e frequentou a Cornell University, primeiro como aluna de graduação e depois como estudante de pós-graduação. Após receber seu Ph.D., McClintock permaneceu em Cornell por vários anos, colaborando com um grupo ilustre de geneticistas de milho, incluindo George Beadle, R. A. Emerson, Charles Burnham, Marcus Rhoades e Lowell Randolph. Juntos, esses pesquisadores desenvolveram os materiais e métodos da genética do milho em uma rica disciplina intelectual. McClintock tinha um papel importante neste trabalho. Devido a sua habilidade na análise citológica, ela teve sucesso em identificar cada um dos 10 cromossomos do milho e foi capaz de conectá-los com grupos de ligação. Ela também fez um trabalho pioneiro sobre o mecanismo de *crossing-over* e a origem do nucléolo. Entretanto, a conquista mais notável de McClintock foi esclarecer as propriedades genéticas dos elementos de transposição. A maior parte de sua pesquisa sobre este assunto foi feita durante os muitos anos em que ela passou no Genetics Department da Carnegie Institution em Cold Spring Harbor, New York. Lá ela teve liberdade para continuar seus estudos sobre mutação mediada por transposon e quebras cromossômicas.

McClintock publicou seu primeiro relato sobre elementos de transposição em 1948. Seguiram-se vários outros relatos, incluindo uma publicação importante no *Cold Spring Harbor Symposium on Quantitative Biology* em 1951. Por muitos motivos, as idéias defendidas por ela nestas publicações não foram bem recebidas. O conceito de transposição contradizia a visão estabelecida de que os genes ocupam posições fixas nos cromossomos. Os dados de McClintock eram complexos, e ela teve dificuldades em comunicá-los a seus colegas. Além disso, a transposição não parecia ser um fenômeno geral. Embora houvesse pouca dúvida de que ela ocorria em milho, ninguém a tinha visto em outros organismos. Esta situação mudou nos anos 1960 e 1970, quando foi descoberta



Fig. 1 ■ Barbara McClintock.

transposição em bactérias e *Drosophila*. Finalmente, o mundo acordou para o amplo significado das idéias de McClintock.

Ela era altamente respeitada por seus colegas. Em 1944, na idade relativamente jovem de 42 anos, ela foi admitida na prestigiosa National Academy of Sciences of the United States. Em 1945, ela foi eleita presidente da Genetics Society of America e em 1970 recebeu a National Medal of Science. Seu Prêmio Nobel veio em 1983, 35 anos após sua primeira publicação sobre elementos de transposição. No dia em que o prêmio foi anunciado, ela estava na floresta coletando cogumelos. Durante sua vida, McClintock foi um modelo de "individualidade obstinada e disciplina".<sup>1</sup> Ela morreu em 1992, alguns meses após seu aniversário de noventa anos.

<sup>1</sup>Federoff, N. V. 1994. Barbara McClintock (16 de junho de 1902 – 2 de setembro de 1992). *Genetics* 136:1-10.

como esperado, alguns mostravam manchas de pigmento marrom-púrpura (Fig. 18.8). McClintock imaginou que, em tais mosaicos, o alelo inibidor  $C^I$  havia perdido algo durante o desenvolvimento do endosperma, levando a um clone de tecido que era capaz de fazer o pigmento. O genótipo em tal clone seria  $-CC$ , onde o traço indica a falta do alelo  $C^I$ .

O mecanismo que McClintock tinha proposto para explicar a perda do alelo  $C^I$  está diagramado na Fig. 18.9. Uma quebra no sítio indicado pela seta destaca um segmento do cromossomo de seu centrômero, criando um fragmento

acêntrico. Tal fragmento tende a ser perdido durante divisão celular. Assim, todos os descendentes desta célula não terão parte do cromossomo de origem paterna. Como o fragmento perdido leva o alelo  $C^I$ , nenhuma das células neste clone é inibida em formar pigmento. Se algumas delas produzem uma parte de aleurona, aparecerá uma mancha de cor, criando um grão similar ao mostrado na Fig. 18.8.

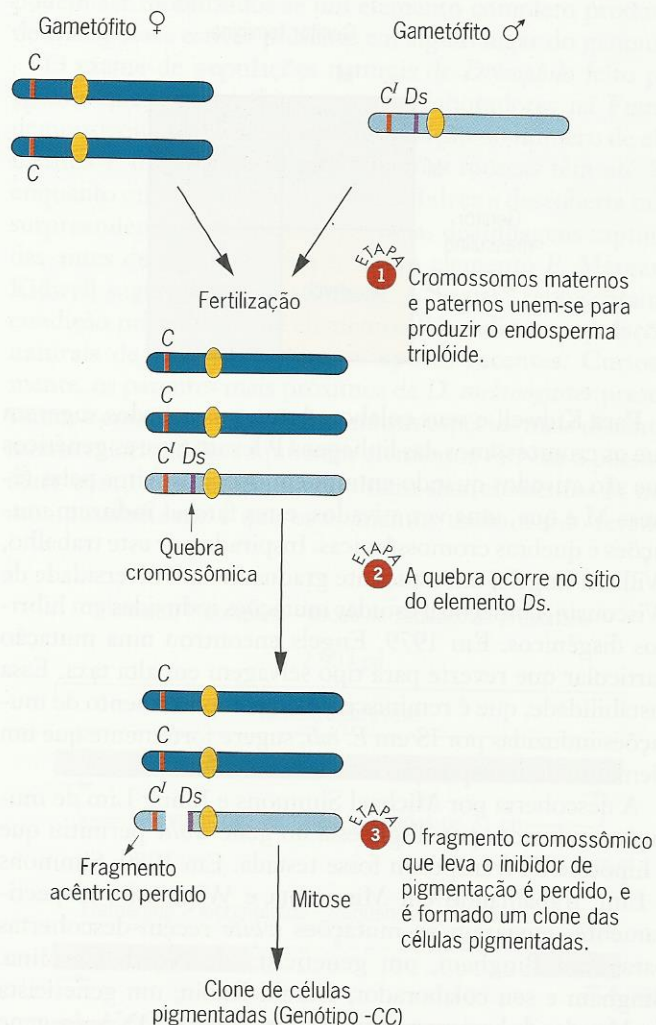
McClintock descobriu que a quebra responsável por estes grãos em mosaico ocorria em um sítio particular no cromossomo 9. Ela chamou o fator que produziu estas quebras de  $Ds$ ,





**Fig. 18.8** ■ Grão de milho (visto por cima) mostrando a perda do alelo  $C'$  para inibição de pigmentação na aleurona. As manchas pigmentadas são -CC, enquanto as manchas incolores são  $C'CC$ .

para **Dissociação**. Entretanto, em si este fator foi incapaz de induzir quebra cromossômica. De fato, McClintock descobriu que  $Ds$  tinha que ser estimulado por outro fator, chamado  $Ac$ , de **Ativador**. O fator  $Ac$  estava presente em alguns estoques de milho mas ausente em outros. Quando foram cruzados estoques diferentes,  $Ac$  podia ser combinado com  $Ds$  para criar a condição que resultou em quebra cromossômica.



**Fig. 18.9** ■ Quebra cromossômica causada pelo elemento de transposição  $Ds$  em milho. O alelo  $C$  no braço curto do cromossomo 9 produz pigmentação normal na aleurona; o alelo  $C'$  inibe essa pigmentação.

Este sistema de dois fatores  $Ac/Ds$  deu uma explicação para a instabilidade genética que McClintock observara no cromossomo 9. Experimentos adicionais demonstraram que esta era apenas uma das muitas instabilidades presentes no genoma de milho. McClintock descobriu outros casos de quebra em sítios diferentes do cromossomo 9 e também em outros cromossomos. Como a quebra nestes sítios depende de ativação por  $Ac$ , ela concluiu que fatores  $Ds$  também estavam envolvidos. Para explicar todas estas observações, McClintock propôs que  $Ds$  poderia existir em muitos sítios diferentes no genoma e que podia mover-se de um sítio para outro.

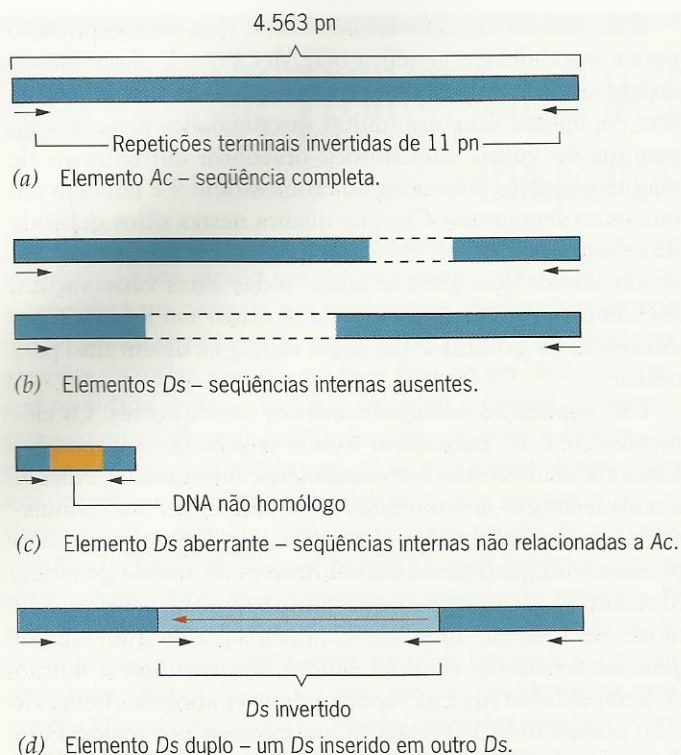
Esta explicação nasceu de análises subsequentes. Os elementos  $Ac$  e  $Ds$  pertencem a uma família de transposons. Estes elementos estão estruturalmente relacionados entre si e podem inserir-se em muitos sítios diferentes nos cromossomos. Várias cópias dos elementos  $Ac$  e  $Ds$  em geral estão presentes no genoma de milho. Através de análise genética, McClintock demonstrou que tanto  $Ac$  quanto  $Ds$  podem mover-se. Quando um destes elementos era inserido em um gene ou próximo a ele, McClintock observou que a função do gene era alterada e às vezes totalmente abolida. Assim,  $Ac$  e  $Ds$  podem induzir mutações inserindo-se nos genes. Para destacar este efeito na expressão gênica, McClintock chamou os transposons  $Ac$  e  $Ds$  de **elementos controladores**.

O seqüenciamento do DNA mostrou que os elementos  $Ac$  consistem em 4.563 pares de nucleotídeos delimitados por repetições invertidas que têm 11 pares de nucleotídeos de tamanho (Fig. 18.10a). Tais repetições terminais invertidas são essenciais para a transposição. Cada elemento  $Ac$  também é flanqueado por repetições diretas com 8 pares de nucleotídeos de tamanho. Como as repetições diretas são criadas no momento em que o elemento é inserido no cromossomo, elas são duplicações dirigidas a sítios, e não partes integrais do elemento.

Ao contrário dos elementos  $Ac$ , os elementos  $Ds$  são estruturalmente heterogêneos. Todos eles possuem as mesmas repetições terminais invertidas dos elementos  $Ac$ , demonstrando que pertencem à mesma família de transposons, mas suas seqüências internas variam. Alguns elementos  $Ds$  parecem ter sido derivados de elementos  $Ac$  pela perda de seqüências internas (Fig. 18.10b). As deleções nestes elementos podem ter sido causadas por síntese incompleta de DNA durante replicação ou transposição. Outros elementos  $Ds$  contêm DNA não- $Ac$  entre suas repetições terminais invertidas (Fig. 18.10c). Estes membros incomuns da família  $Ac/Ds$  são chamados de elementos **aberrantes**  $Ds$ . Uma terceira classe de elementos  $Ds$  é caracterizada por uma disposição peculiar *piggybacking* (do tipo carreta de transporte de automóveis) (Fig. 18.10d); um elemento  $Ds$  é inserido em outro mas na orientação invertida. Estes chamados elementos  $Ds$  *duplos* são aparentemente responsáveis pela quebra de cromossomos que McClintock observou em seus experimentos.

As atividades dos elementos  $Ac/Ds$  – excisão e transposição e todos os seus correlatos genéticos, incluindo mutação e quebra de cromossomos – são causadas por uma transposase





**Fig. 18.10** ■ Organização estrutural dos membros da família Ac/Ds de elementos de transposição em milho. São indicadas as repetições terminais invertidas (setas curtas abaixo) e os tamanhos das seqüências de DNA (em pares de nucleotídeos, pn).

codificada pelos elementos *Ac*. A transposase *Ac* aparentemente interage com seqüências nas pontas de elementos *Ac* e *Ds* ou próximo a elas, catalisando seu movimento. Deleções ou mutações no gene que codifica a transposase abolem esta função catalítica. Assim, os elementos *Ds*, que têm tais lesões, não podem ativar a si próprios. Entretanto, eles podem ser ativados se um elemento *Ac* produtor de transposase estiver presente em algum lugar no genoma. A transposase feita por este elemento pode difundir-se pelo núcleo, ligar-se ao elemento *Ds* e ativá-lo. A transposase *Ac* é, portanto, uma proteína de ação *trans*.

Transposons relacionados com os elementos *Ac/Ds* foram encontrados em outras espécies, incluindo animais. Talvez o mais bem estudado destes elementos seja um chamado *hobo*, assim denominado por sua capacidade de se transpor. O elemento *hobo* é encontrado em algumas espécies de *Drosophila*.

### Elementos *P* e Disgenesia Híbrida em *Drosophila*

Algumas das mais amplas pesquisas em elementos de transposição enfocaram os elementos *P* de *Drosophila melanogaster*. Esses transposons foram identificados pela cooperação de geneticistas que trabalhavam em diversos laboratórios diferentes. Em 1977, Margaret e James Kidwell, trabalhando

em Rhode Island, e John Sved, trabalhando na Austrália, descobriram que cruzamentos entre algumas linhagens de *Drosophila* produzem híbridos com uma distribuição de características aberrantes, incluindo mutação freqüente, quebra cromossômica e esterilidade. O termo **disgenesia híbrida**, derivado da raiz grega que significa “uma deterioração de qualidade”, foi usado para indicar esta síndrome de anomalias.

Kidwell e seus colaboradores descobriram que podiam classificar as linhagens de *Drosophila* em dois tipos principais com base na propriedade de elas produzirem ou não híbridos disgênicos em cruzamentos testes. Os dois tipos de linhagens foram designados M e P. Apenas cruzamentos entre as linhagens M e P produziam híbridos disgênicos, e eles o fazem apenas se o macho no cruzamento for da linhagem P. Cruzamentos entre duas linhagens P diferentes, ou entre duas linhagens M diferentes, produzem híbridos que são normais. Podemos resumir os fenótipos da prole híbrida desses cruzamentos diferentes em um quadro simples:

|                   |   | Genitor feminino |        |
|-------------------|---|------------------|--------|
|                   |   | M                | P      |
| Genitor masculino | M | normal           | normal |
|                   | P | disgênico        | normal |

Para Kidwell e seus colaboradores, tais achados sugerem que os cromossomos das linhagens P levam fatores genéticos que são ativados quando entram em ovócitos feitos pelas fêmeas M e que, uma vez ativados, estes fatores induzem mutações e quebras cromossômicas. Inspirado por este trabalho, William Engels, um estudante graduado na Universidade de Wisconsin, começou a estudar mutações induzidas em híbridos disgênicos. Em 1979, Engels encontrou uma mutação particular que reverte para tipo selvagem em alta taxa. Essa instabilidade, que é remanescente do comportamento de mutações induzidas por IS em *E. coli*, sugere fortemente que um elemento de transposição estava envolvido.

A descoberta por Michael Simmons e John Lim de mutações induzidas por disgenesia no gene *white* permitiu que a hipótese de transposon fosse testada. Em 1980, Simmons e Lim, trabalhando em Minnesota e Wisconsin, respectivamente, enviaram as mutações *white* recém-descobertas para Paul Bingham, um geneticista de North Carolina. Bingham e seu colaborador, Gerald Rubin, um geneticista de Maryland, haviam terminado de isolar o DNA do gene *white*. Usando este DNA como uma sonda, Bingham e Rubin foram capazes de isolar DNA dos alelos *white* mutantes e compará-lo ao DNA *white* tipo selvagem. Em cada mutação, eles descobriram que um pequeno elemento havia sido



inserido na região codificante do gene *white*. Experimentos adicionais demonstraram que tais elementos estão presentes em várias cópias e em locais diferentes nos genomas das linhagens P. Entretanto, eles estão completamente ausentes dos genomas das linhagens M. Os geneticistas começaram portanto a chamar estes transposons específicos de linhagens P de **elementos P**.

Análise da sequência de DNA mostrou que os elementos P variam de tamanho. Os maiores elementos têm 2.907 pares de nucleotídeos de tamanho, incluindo repetições terminais invertidas de 31 pares de nucleotídeos. Esses elementos P completos levam um gene que codifica uma transposase. Quando a transposase P se liga perto das pontas de um elemento P completo, ela pode mover esse elemento para um novo local no genoma. Elementos P incompletos (Fig. 18.11) não têm a capacidade de produzir a transposase porque algumas de suas sequências internas estão deletadas. Entretanto, eles possuem as sequências terminais e subterminais que ligam a transposase. Conseqüentemente, tais elementos podem ser mobilizados se um elemento completo produtor de transposase estiver presente em algum lugar do genoma.

O exame de populações naturais de *Drosophila* feito por Dominique Anxolabéhère e seus colaboradores na França demonstrou que há considerável variação no número de elementos P nos cromossomos. Algumas moscas têm até 50, enquanto outras têm apenas alguns. Talvez a descoberta mais surpreendente é que moscas derivadas de linhagens capturadas antes de 1950 não têm nenhum elemento P. Margaret Kidwell sugeriu que essas linhagens “vazias” representam a condição primitiva e que elementos P invadiram populações naturais de *Drosophila* durante épocas recentes. Curiosamente, os parentes mais próximos de *D. melanogaster* preservaram a condição “vazia”, mas outras espécies mais distantemente relacionadas adquiriram elementos P. Não é possível dizer como estas espécies adquiriram seus elementos P, mas uma possibilidade é que os elementos tenham sido levados

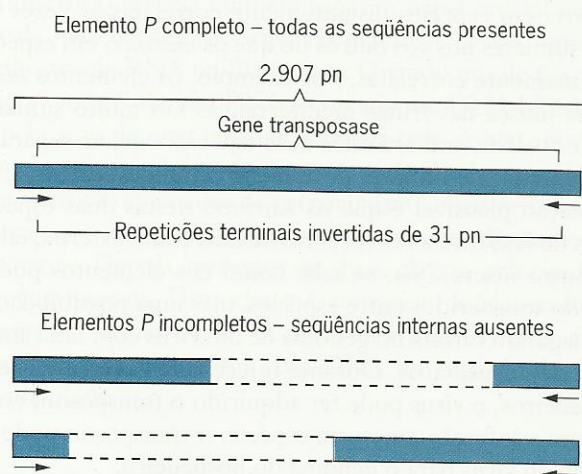
para o genoma por vírus que naturalmente infestam *Drosophila*. Tal processo seria análogo à transdução de *E. coli* por um bacteriófago que leve um elemento IS.

Populações de *Drosophila* que possuem elementos P desenvolveram mecanismos para regular seu movimento. Em algumas linhagens, essa regulação depende do **citotipo**, uma condição celular que é transmitida maternamente através do citoplasma do ovócito. O citotipo P reprime o movimento do elemento P, e o citotipo M o permite. O citotipo P é característico de linhagens que têm elementos P em seus cromossomos, enquanto o citotipo M é característico de linhagens que não têm os elementos P. Quando elementos P são combinados com o citotipo M por meio de cruzamentos apropriados, eles são induzidos a serem transpostos. Embora o papel regulador do citotipo tenha sido estudado por vários pesquisadores, sua natureza molecular ainda é desconhecida.

A transmissão materna do citotipo pode ser vista na prole de cruzamentos recíprocos entre linhagens com citotipos P e M (Fig. 18.12). Existem dois cruzamentos: (1) fêmea com citotipo P × macho com citotipo M e (2) macho com citotipo P × fêmea com citotipo M. A prole de ambos os cruzamentos herda elementos P em seus cromossomos e é, de fato, genotipicamente idêntica. Entretanto, apenas indivíduos do segundo cruzamento permitem movimento de P. Esta diferença entre os cruzamentos recíprocos indica que a condição que permite ou reprime o movimento de P é transmitida apenas pelo genitor feminino, supostamente no citoplasma do ovócito. No cruzamento 1, as fêmeas transmitem o citotipo P para sua prole, que então reprime o movimento de P, enquanto no cruzamento 2 as fêmeas transmitem o citotipo M, que permite o movimento de P.

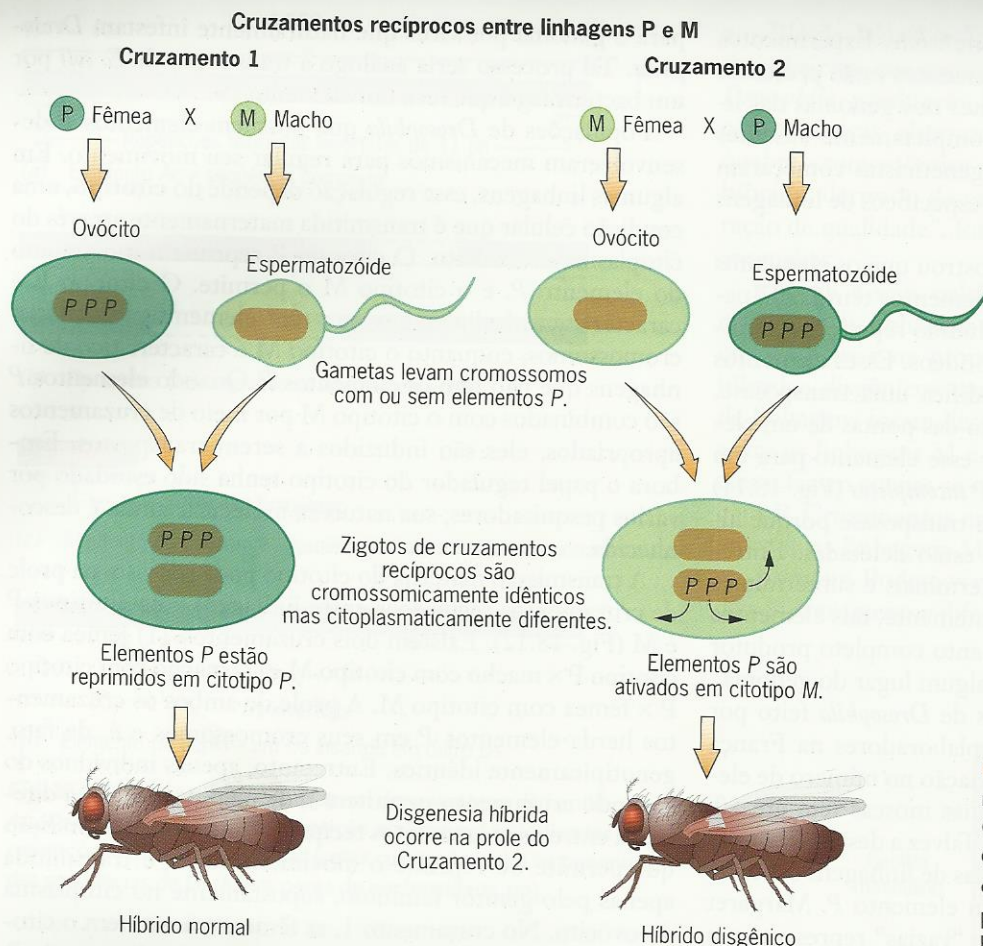
Considerando o dano que pode ser causado pelo amplo movimento do elemento P, parece surpreendente que cruzamentos macho P × fêmea M produzam qualquer prole viável. Donald Rio, Frank Laski e Gerald Rubin mostraram que essas proles são saudáveis porque os elementos P se movem apenas na linhagem germinativa. Nos tecidos somáticos, onde o movimento do elemento P causaria graves problemas, há pouca, se alguma, transposição porque a transposase P não é sintetizada lá. O bloqueio metabólico ocorre no nível da recomposição do RNA; um dos íntrons permanece no transcrito do gene de transposase, criando um códon de fim que termina prematuramente a tradução. Assim, em vez de fazer a transposase, as células somáticas sintetizam uma proteína mais curta que não tem a função catalítica da transposase. Sem esta função, os elementos P são incapazes de se mover, e a mosca é protegida de intenso dano em seus tecidos somáticos.

Geneticistas usam rotineiramente disgenesia híbrida para obter mutações de inserção de P no laboratório. Híbridos disgênicos produzidos cruzando-se machos P com fêmeas M são cruzados para recuperar mutações que ocorreram em suas linhagens germinativas. Tais mutações são detectadas por técnicas apropriadas, tais como o esquema de cruzamento *CIB* de Muller para identificar letais ligados ao X



**Fig. 18.11** ■ Estrutura dos elementos P em *Drosophila* mostrando orientações e tamanhos (em pares de nucleotídeos, pn) de sequências de DNA.





**Fig. 18.12 ■** Disgenesia híbrida mediada por elemento P em *Drosophila*. O citotipo é herdado maternamente. O citotipo P reprime o movimento do elemento P, enquanto o citotipo M o permite. Ocorre disgenesia apenas nos híbridos do citotipo M do cruzamento 2.

(Cap. 13). A disgenesia híbrida tem uma vantagem em relação aos métodos tradicionais de induzir mutações, porque um gene que foi mutado pela inserção de um elemento de transposição é “marcado” com uma sequência conhecida de DNA. A marca de transposon pode subsequentemente ser usada para isolar o gene de uma grande mistura heterogênea de DNA. A mutagenese por **marcação de transposon** é portanto uma técnica padrão em genética molecular. Um Marco na Genética: Transformação de *Drosophila* com Elementos P discute outro emprego importante de elementos P em pesquisa genética.

### Mariner, um Antigo e Generalizado Transposon

Dois dos parentes mais próximos de *Drosophila melanogaster*, *D. simulans* e *D. mauritiana*, possuem um pequeno transposon chamado *mariner* (1.286 pares de nucleotídeos de tamanho com 28 pn de repetições terminais invertidas). Embora *mariner* não esteja presente em *D. melanogaster*, transposons similares estão presentes em muitos outros insetos, nematódos, fungos e mesmo seres humanos. Essa ampla distribuição sugere que os elementos *mariner* são antigos, datando das primeiras épocas evolutivas.

Análise de sequências de *mariners* de muitos organismos diferentes sugeriu que tais elementos são algumas vezes transferidos horizontalmente entre espécies. Hugh Robertson mostrou que elementos do tipo *mariner* estão presentes em várias ordens diferentes de insetos. Curiosamente, os *mariners* em espécies distantemente correlatas às vezes são mais similares uns aos outros do que os *mariners* em espécies proximamente correlatas. Por exemplo, os elementos *mariner* na mosca-das-frutas mediterrâneas são muito similares aos de abelhas, embora estas duas espécies estejam separadas por centenas de milhões de anos no tempo evolutivo. Uma explicação plausível é que os *mariners* nestas duas espécies sejam de fato invasores recentes de uma fonte externa, talvez um outro inseto. Não se sabe como tais elementos podem ter sido transferidos entre espécies, mas uma possibilidade é que pegaram carona no genoma de um vírus com uma ampla gama de hospedeiros. Durante infecção em uma espécie de hospedeiros, o vírus pode ter adquirido o transposon; então durante a infecção em outra espécie, o transposon pode ter pulado do vírus para o genoma do hospedeiro.

Seqüenciamento do DNA também identificou outro grupo de transposons relacionado com os elementos *mariner*. Eles são um pouco maiores, 1,6 a 1,7 quilobases de



tamanho, com repetições invertidas em suas pontas. Suas seqüências repetidas variam tipicamente de 54 a 234 pares de nucleotídeos. O transposon mais bem estudado neste grupo, chamado de *Tc1*, é encontrado no nematódeo *Caenorhabditis elegans*. Outros elementos do tipo *Tc1* foram encontrados em vertebrados. Os elementos *Tc1* e *mariner* pertencem portanto a uma superfamília de transposons de origem antiga.

### PONTOS IMPORTANTES

- O elemento de transposição *Ds* de milho, descoberto devido a sua capacidade de quebrar cromossomos, é ativado por outro elemento de transposição, *Ac*, que codifica uma transposase.
- Os elementos de transposição *P* são responsáveis por disgenesia híbrida, uma síndrome de anomalias da linhagem germinativa que ocorre na prole de cruzamentos entre as linhagens *P* e *M* de *Drosophila*.
- Os transposons *mariner*, que foram originalmente descobertos em algumas espécies de *Drosophila*, são amplamente distribuídos no reino animal e parecem ter-se espalhado horizontalmente entre espécies diferentes.

## RETROVÍRUS E RETROTRANSPOSONS

Retrovírus e elementos de transposição correlatos usam a enzima transcriptase reversa para copiar RNA em DNA. As cópias de DNA destas entidades são subsequenteiramente inseridas em posições diferentes no DNA genômico.

Além dos transposons de recortar e colar tais como *Ac*, *P* e *mariner*, os genomas eucarióticos contêm elementos de transposição cujos movimentos dependem de transcrição reversa de RNA em DNA. Essa reversão no fluxo de informação genética levou os geneticistas a chamar tais elementos de **retrotransposons**, do prefixo latino que indica “reverso”. A transcrição reversa também tem um papel crucial nos ciclos de vida de alguns vírus. Os genomas desses vírus são compostos de RNA unifilamentar. Quando um de tais vírus infecta uma célula, seu RNA é copiado em DNA bifilamentar pela ação de uma enzima chamada transcriptase reversa. Como a informação genética se move do RNA para o DNA, esses vírus são chamados de **retrovírus**. Começaremos nossa investigação dos retrotransposons com uma discussão dos retrovírus. Depois, veremos as duas classes principais de retrotransposons. Uma classe é composta de elementos similares às formas de DNA dos retrovírus. A outra classe é composta de elementos cujo DNA foi copiado de RNA poliadenilado.

### Retrovírus

Os retrovírus foram descobertos por meio de estudo das causas de alguns tipos de tumores em galinhas, gatos e ca-

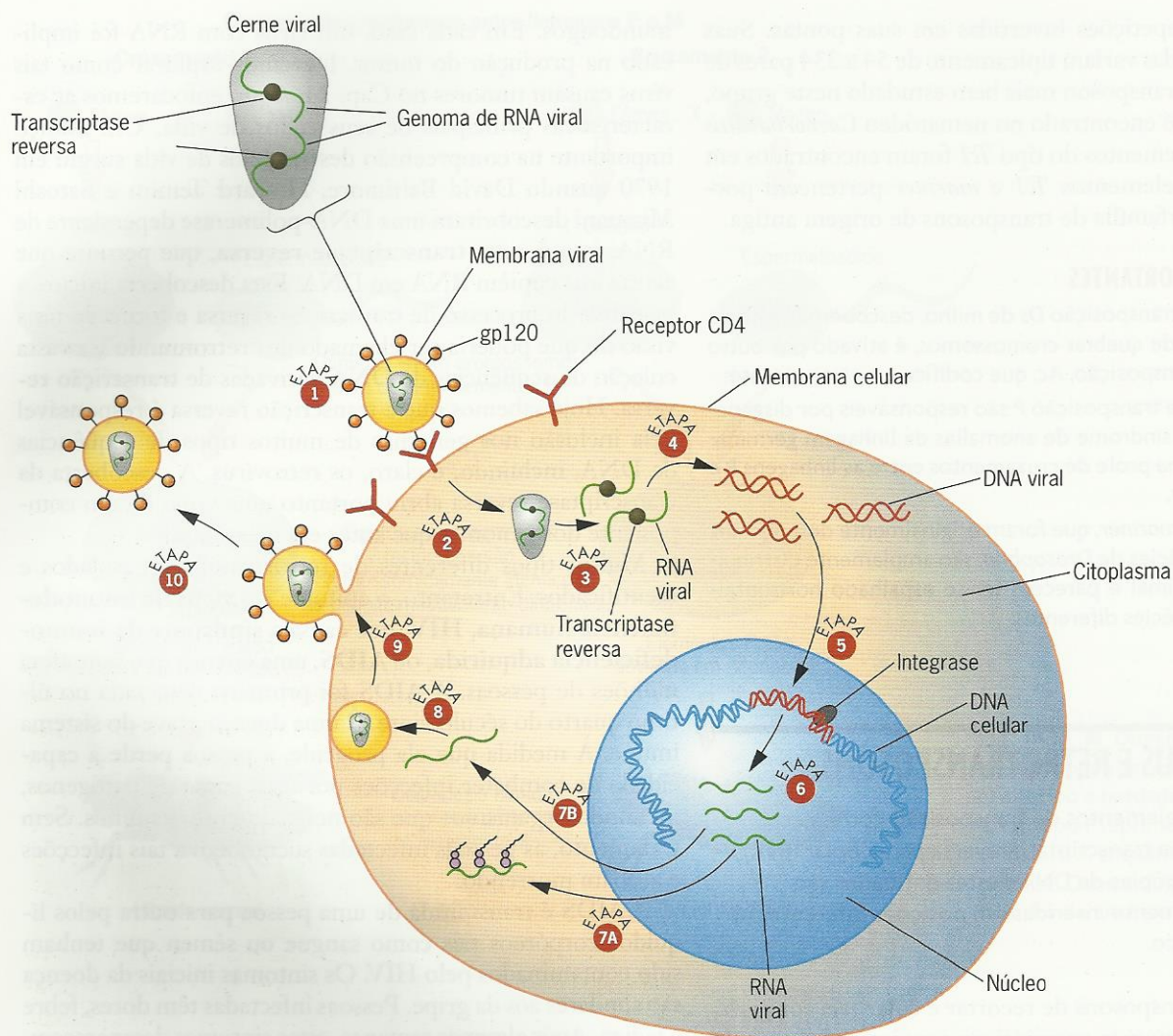
mundongos. Em cada caso, um vírus com RNA foi implicado na produção do tumor. Devemos explorar como tais vírus causam tumores no Cap. 24. Aqui, enfocaremos as características principais de seus ciclos de vida. Um avanço importante na compreensão destes ciclos de vida surgiu em 1970 quando David Baltimore, Howard Temim e Satoshi Mizutani descobriram uma DNA-polimerase dependente de RNA, isto é, uma **transcriptase reversa**, que permite que estes vírus copiem RNA em DNA. Esta descoberta iniciou a pesquisa do processo de transcrição reversa e forneceu uma visão do que poderia ser chamado de “retromundo”, a vasta coleção de seqüências de DNA derivadas de transcrição reversa. Hoje sabemos que a transcrição reversa é responsável pela inclusão nos genomas de muitos tipos de seqüências de DNA, incluindo, é claro, os retrovírus. A descoberta da transcriptase reversa abriu portanto uma visão de um componente dos genomas que antes era inexplorado.

Muitos tipos diferentes de retrovírus foram isolados e identificados. Entretanto, o epítome é o **vírus da imunodeficiência humana, HIV**, que causa a **síndrome da imunodeficiência adquirida**, ou **AIDS**, uma doença que hoje afeta milhões de pessoas. A AIDS foi primeiro detectada no último quarto do século vinte. É uma doença grave do sistema imune. À medida que ela progride, a pessoa perde a capacidade de combater infecções por uma gama de patógenos, incluindo organismos que são normalmente benignos. Sem tratamento, as pessoas infectadas sucumbem a tais infecções e acabam morrendo.

A AIDS é transmitida de uma pessoa para outra pelos líquidos corpóreos tais como sangue ou sêmen que tenham sido contaminados pelo HIV. Os sintomas iniciais da doença são similares aos da gripe. Pessoas infectadas têm dores, febre e fadiga. Após algumas semanas, estes sintomas desaparecem, e a saúde parece ter sido restaurada. Este estado assintomático pode durar vários anos. Entretanto, o vírus continua a se multiplicar e se espalhar pelo corpo, tendo como alvo células especializadas com papéis importantes no sistema imunológico. Ao final, essas células estão de tal modo depletadas pela ação mortal do vírus que o sistema imunológico falha e patógenos oportunistas se instalam. Muitos tipos de doenças, tais como pneumonia, podem ocorrer. A AIDS é uma causa importante de morte entre subpopulações de muitos países – por exemplo, entre usuários de drogas intravenosas e trabalhadores da indústria do sexo – e, na África subsaariana, é uma causa importante de morte na população como um todo.

Devido a sua letalidade e condição pandêmica, HIV/AIDS tem sido o foco de uma grande quantidade de pesquisas. Um resultado deste esforço foi uma compreensão detalhada do ciclo de vida do HIV (Fig. 18.13). O vírus quase esférico entra na célula hospedeira interagindo com proteínas receptoras específicas, chamadas receptores CD4, que estão localizadas na superfície da célula. Essa interação é mediada por uma glicoproteína (uma proteína à qual se ligam açúcares) chamada gp120, que fica inserida na membrana lipídica que circunda a partícula viral. Uma vez que gp120 se tenha ligado ao receptor CD4, as membranas viral e celular se fun-





**Fig. 18.13** ■ O ciclo de vida do HIV.

- ETAPA 1. HIV ancora-se à célula alvo por uma interação entre a proteína viral gp120 e a proteína receptora CD4 celular.
- ETAPA 2. As membranas celular e viral fundem-se, permitindo que o cerne viral entre na célula.
- ETAPA 3. RNA e proteínas associadas são liberados do cerne viral.
- ETAPA 4. A transcriptase reversa catalisa a síntese do DNA viral bifilamentar a partir do RNA viral unifilamentar no citoplasma.
- ETAPA 5. A integrase catalisa a inserção de DNA viral no DNA celular no núcleo.
- ETAPA 6. A RNA-polimerase celular transcreve DNA viral em RNA viral.
- ETAPA 7A. Parte do RNA viral serve como mRNA para a síntese das proteínas virais.
- ETAPA 7B. Parte do RNA viral forma os genomas dos vírus da prole.
- ETAPA 8. Partículas de vírus da prole são montadas perto da membrana celular.
- ETAPA 9. Prole das partículas virais são removidas da célula por brotamento.
- ETAPA 10. Prole das partículas de vírus estão livres para infectar outras células.

dem e a partícula viral entra na célula. Dentro da célula, a membrana lipídica e a capa de proteína que circunda a partícula viral são removidas, e os materiais dentro do cerne do vírus são liberados no citoplasma da célula. Esse cerne contém duas moléculas de RNA unifilamentar idênticas, o genoma do vírus, e um pequeno número de proteínas que facilitam a replicação do genoma, incluindo duas moléculas da transcriptase reversa viral, uma ligada a cada filamento do RNA viral.

A transcriptase reversa do vírus, e outras transcriptases reversas, converte o RNA unifilamentar em DNA bifilamentar. As moléculas bifilamentares de DNA resultantes são então inseridas em posições aleatórias nos cromossomos da célula infectada, populando o genoma das células com muitas cópias do genoma viral. Essas cópias podem ser então transcritas pela RNA-polimerase da célula para produzir uma grande quantidade de RNA viral, que serve para dirigir a síntese de proteínas virais e também fornece RNA



genômico para a montagem de novas partículas virais. Tais partículas são expulsas da célula por um processo de brotamento na membrana da célula. As partículas expulsas podem então infectar outras células interagindo com os receptores CD4 de suas superfícies. Deste modo, o material genético do HIV é replicado e disseminado por uma população de células imunológicas suscetíveis.

O genoma do HIV, com um pouco mais de 10 quilobases de tamanho, contém vários genes. Três desses genes, chamados *gag*, *pol* e *env*, são encontrados em todos os outros retrovírus. O gene *gag* codifica proteínas da partícula viral, o gene *pol* codifica a transcriptase reversa e outra enzima chamada integrase, que catalisa a inserção do DNA do genoma de HIV nos cromossomos de uma célula hospedeira, e o gene *env* codifica as glicoproteínas inseridas no envoltório lipídico do vírus.

Examinemos melhor a replicação do genoma do HIV (Fig. 18.14). Este processo, catalisado pela transcriptase reversa, começa com a síntese de um só filamento de DNA complementar ao único filamento de RNA do genoma viral. Ele é iniciado por um tRNA que é complementar a uma sequência chamada PBS (sítio de ligação do *primer*) situada à esquerda do centro no RNA do HIV (etapa 1 na Fig. 18.14). Esse tRNA é embalado pré-hibridizado a PBS no cerne do HIV. Após a transcriptase reversa catalisar a síntese da ponta 3' do DNA viral, a ribonuclease H (RNase H) degrada o RNA genômico no dúplice RNA-DNA (etapa 2). Essa degradação deixa a sequência repetida (R) do filamento de DNA nascente livre para hibridizar com a sequência R na ponta 3' do RNA do HIV. O resultado final é que a região R do filamento de DNA nascente "salta" da ponta 5' do RNA do HIV para a ponta 3' do RNA do HIV (etapa 3). A transcriptase reversa em seguida amplia a cópia de DNA usando a região 5' do RNA do HIV como molde (etapa 4).

Na etapa 5, a RNaseH degrada todo o RNA no dúplice RNA-DNA exceto uma pequena região, a via de polipurinas, que é composta principalmente das purinas adenina e guanina. Esta via de polipurinas é usada para iniciar a síntese do segundo filamento de DNA da fração 3' do genoma do HIV (etapa 6). Após o tRNA e o RNA genômico presentes nos dúplices RNA-DNA serem removidos (etapa 7), ocorre um segundo "salto" de DNA durante o qual o PBS na ponta 5' do segundo filamento de DNA se hibridiza com o PBS complementar na ponta 5' do primeiro filamento de DNA (etapa 8). As pontas hidroxila 3' dos dois filamentos de DNA são então usadas para iniciar a síntese de DNA de modo a completar a síntese do DNA bifilamentar do HIV (etapa 9). Note que a conversão do RNA viral em DNA viral produz sequências marcantes em ambas as pontas da molécula de DNA. Estas sequências, chamadas **longas repetições terminais (LTR)**, são necessárias para integração do genoma viral ao DNA da célula hospedeira.

A integração (Fig. 18.15) do DNA viral é catalisada pela enzima integrase, que tem atividade de endonuclease. A integrase primeiro produz pontas 3' recuadas no DNA do HIV por cortes unifilamentares perto das pontas de ambas as LTR

(etapa 1). Tais pontas recuadas são em seguida usadas para a ligação catalisada por integrase nas ligações fosfodiéster em uma sequência alvo no DNA da célula hospedeira. Este processo resulta na formação de novas ligações fosfodiéster entre as pontas 3' do DNA do HIV e 5' fosfato do DNA hospedeiro (etapa 2). No estágio final de integração, as enzimas de reparo do DNA da célula hospedeira preenchem os espaços unifilamentares para produzir um genoma de DNA do HIV inserido covalentemente no DNA cromossômico da célula hospedeira (etapa 3). Note que a sequência alvo no sítio de integração é duplicada no processo. O genoma integrado do HIV torna-se portanto parte permanente do genoma da célula hospedeira, replicando-se como qualquer outro segmento do DNA hospedeiro.

Retrovírus integrados de muitos tipos diferentes estão presentes em genomas de vertebrados, incluindo o nosso. Como tais retrovírus são replicados juntamente com o resto do DNA, eles são transmitidos para as células filhas durante divisão e, se estiverem integrados às células da linhagem germinativa, eles também são transmitidos para a geração seguinte pelos gametas. Geneticistas chamam as sequências herdáveis de DNA que são derivadas de transcrição reversa e integração dos genomas virais de **retrovírus endógenos**. Na maior parte, estas sequências perderam sua capacidade de produzir partículas virais infecciosas. Eles são, portanto, vestígios inócuos de infecções virais muito antigas. O HIV não é um retrovírus endógeno, mas, se ele deve perder seu potencial letal e ser transmitido sob forma integrada através da linhagem germinativa, ele pode tornar-se um.

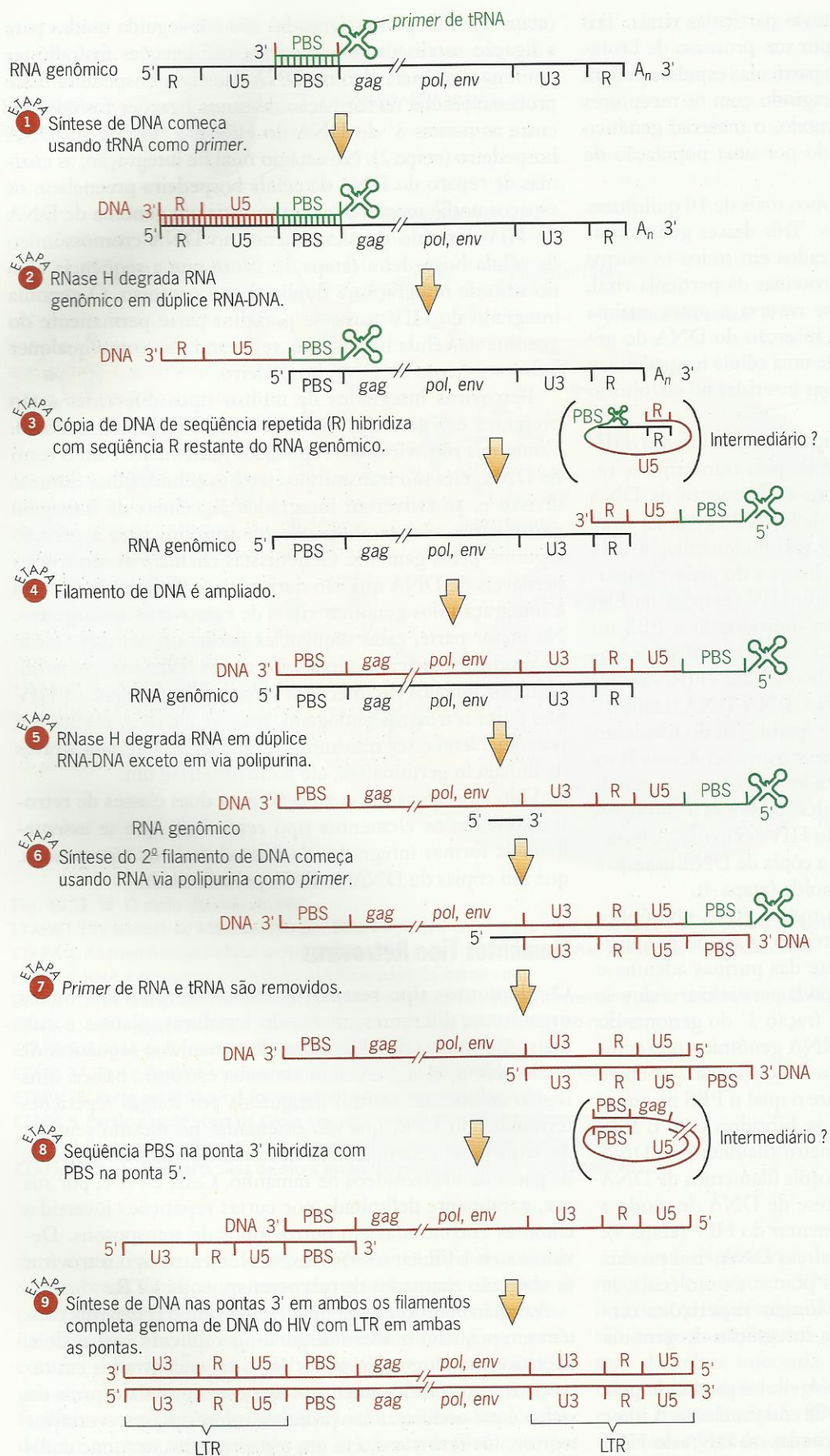
Voltemos agora nossa atenção para duas classes de retrotransposons: os elementos tipo retrovírus, que se assemelham às formas integradas de retrovírus, e os transposons, que são cópias de DNA de RNA poliadenilado.

## Elementos Tipo Retrovírus

Os **elementos tipo retrovírus** são encontrados em muitos organismos diferentes, incluindo leveduras, plantas e animais. A despeito das diferenças de tamanho e sequência de nucleotídeos, eles todos têm a mesma estrutura básica: uma região codificante central flanqueada por longas repetições terminais, ou LTR, que são orientadas no mesmo sentido. As sequências repetidas têm tipicamente algumas centenas de pares de nucleotídeos de tamanho. Cada LTR é, por sua vez, geralmente delimitada por curtas repetições invertidas como as encontradas em outros tipos de transposons. Devido a suas LTR características, os elementos tipo retrovírus às vezes são chamados de **retrotransposons LTR**.

A região codificante de um elemento tipo retrovírus contém um pequeno número de genes, geralmente apenas dois. Eles são homólogos aos genes *gag* e *pol* encontrados em retrovírus; *gag* codifica uma proteína estrutural da cápsula do vírus, e *pol* codifica uma proteína transcriptase reversa/integrase. Os retrovírus têm um terceiro gene, *env*, que codifica um componente protéico do envoltório do vírus. Nos





**Fig. 18.14 ■** Conversão do genoma de RNA unifilar do HIV embalado no vírion na forma de DNA bifilar que é integrado ao DNA da célula hospedeira. A transcriptase reversa viral usa o RNA genômico como molde para sintetizar um filamento complementar de DNA. A transcriptase reversa e o primer de tRNA são ambos embalados com o RNA do HIV no cerne do vírion. R, sequência repetida; U5, sequência única perto da ponta 5'; U3, sequência única perto da ponta 3'; PBS, sítio de ligação do primer; cauda A<sub>n</sub> poli (A); gag, pol e env, sequências que codificam proteínas do HIV, via de polipurina, sequência de nucleotídeos rica nas purinas adenina e guanina; LTR, longa repetição terminal.







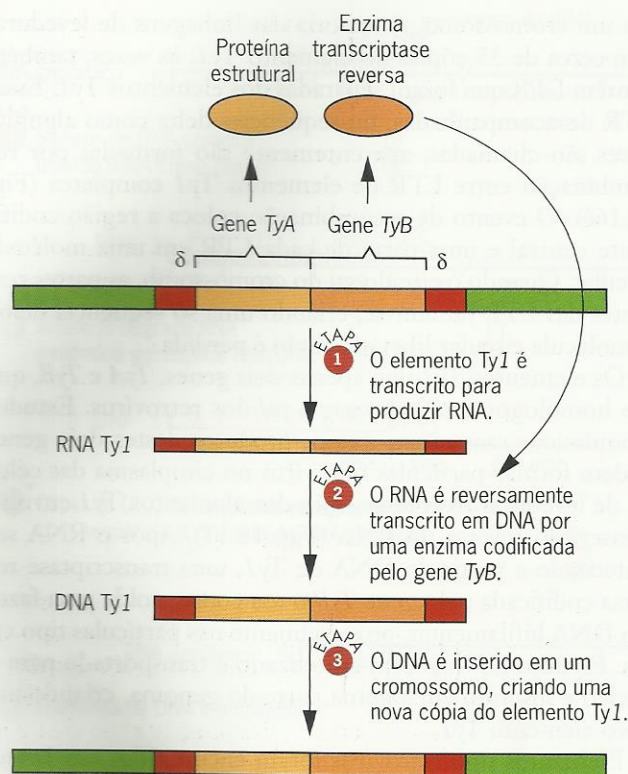


Fig. 18.17 ■ Transposição do elemento Ty1 de levedura.

porque contém um gene similar ao gene *env* dos retrovírus. Tanto elementos *copia* quanto *gypsy* formam partículas tipo vírus dentro das células de *Drosophila*. Entretanto, apenas as partículas que contêm RNA *gypsy* podem mover-se através das membranas celulares, possivelmente porque também contêm o produto do gene *gypsy env*. O elemento *gypsy* parece, portanto, ser um retrovírus genuíno. Muitas outras famílias de transposons tipo retrovírus foram encontradas em *Drosophila*, mas suas atividades são pouco compreendidas.

## Retroposons

**Retroposons**, ou retrotransposons não-LTR, são uma grande e amplamente distribuída classe de retrotransposons, incluindo os elementos *F*, *G* e *I* de *Drosophila* e vários tipos de elementos em mamíferos. Tais elementos movem-se por uma molécula de RNA que é reversamente transcrita em DNA, provavelmente por uma proteína codificada pelos próprios elementos. Embora criem uma duplicação de sítio alvo quando se inserem em um cromossomo, eles não têm repetições invertidas ou diretas como partes integrais de seus terminos. Ao contrário, eles são distinguidos por uma sequência homogênea de pares de bases A:T em uma ponta. Esta sequência é derivada de transcrição reversa da cauda poli(A) que é adicionada perto da ponta 3' do RNA retroposon durante sua maturação. Retroposons integrados apresentam, portanto, um vestígio de sua origem como transcritos reversos de RNA poliadenilados.

Em *Drosophila*, retroposons especiais são encontrados nas pontas (telômeros) dos cromossomos, onde desempenham a função crítica de repor o DNA que foi perdido por replicação cromossômica incompleta. A cada rodada de replicação do DNA, um cromossomo torna-se mais curto. O encurtamento ocorre porque a DNA-polimerase só pode mover-se em um sentido, adicionando nucleotídeos à ponta 3' de um *primer* (Cap. 10). Geralmente, o *primer* é o RNA, e, quando ele é removido, uma região unifilamentar é deixada na ponta do DNA bifilamentar. Na próxima rodada de replicação, o filamento deficiente produz uma dupla hélice que é mais curta que a original. À medida que este processo continua, ciclo após ciclo, os cromossomos perdem material de sua ponta.

Para contrabalançar essa perda, *Drosophila* desenvolveu um mecanismo curioso envolvendo pelo menos dois retroposons diferentes, um chamado de *HeT-A* e outro de *TART* (retrotransposon associado a telômero). Mary Lou Pardue, Robert Levis, Harald Biessmann, James Mason e seus colaboradores mostraram que estes dois elementos preferencialmente são transpostos para as pontas dos cromossomos, ampliando-os em várias quilobases. Ao final, as sequências transpostas são perdidas por replicação incompleta do DNA, mas então ocorre uma nova transposição para restaurá-las. Os retroposons *HeT-A* e *TART* desempenham portanto a função importante de regenerar as pontas perdidas dos cromossomos.

## PONTOS IMPORTANTES

- Os genomas dos retrovírus são compostos de RNA unifilamentar contendo pelo menos três genes: *gag* (codificando proteínas estruturais da partícula viral), *pol* (codificando uma proteína transcriptase reversa/integrase) e *env* (codificando uma proteína inserida no envoltório lipídico do vírus).
- O retrovírus HIV humano infecta células do sistema imunológico e causa a doença AIDS ameaçadora à vida.
- Os elementos tipo retrovírus possuem genes homólogos a *gag* e *pol*, mas não a *env*.
- Elementos tipo retrovírus e o DNA de retrovírus inseridos em cromossomos celulares são demarcados por sequências de longas repetições terminais (LTR).
- Os retroposons não têm LTR; entretanto, em uma ponta eles têm uma sequência de pares de bases A:T derivada da transcrição reversa de uma cauda poli(A) ligada ao RNA do retroposon.
- Os retroposons *HeT-A* e *TART* são componentes das pontas dos cromossomos de *Drosophila*.

## ELEMENTOS DE TRANSPOSIÇÃO EM HUMANOS

O genoma humano contém uma diversa gama de elementos de transposição que coletivamente contribuem com 44% de todo o DNA humano.

Com o sequenciamento do genoma humano, hoje é possível avaliar o significado dos elementos de transposição



em nossa espécie. Pelo menos 44% do DNA humano é derivado de elementos de transposição, incluindo elementos tipo retrovírus (8% do genoma sequenciado), retroposons (33%) e várias famílias de transposons de DNA que são transpostos por um mecanismo de cortar e colar (3%).

O elemento de transposição dominante é um retroposon chamado de **L1**. Este elemento pertence a uma classe de seqüências conhecidas como **longos elementos nucleares intercalares**, ou **LINE**. Os elementos **L1** completos têm cerca de 6 kb de tamanho, têm um promotor interno que é reconhecido pela RNA-polimerase II e têm duas matrizes de leitura abertas, ORF1, que codifica uma proteína de ligação a ácido nucleico, e ORF2, que codifica uma proteína com atividades de endonuclease e transcriptase reversa. O genoma humano contém entre 3.000 e 5.000 elementos **L1** completos. Além disso, ele contém mais de 500.000 elementos **L1** que são truncados em suas pontas 5'. Esses elementos **L1** incompletos são transposicionalmente inativos. Cada elemento **L1** no genoma, seja completo ou incompleto, geralmente é flanqueado por uma curta duplicação de sítio alvo.

Os elementos **L1** são autênticos retroposons. Sua transposição envolve a transcrição de um elemento **L1** completo em RNA e a transcrição reversa deste RNA em DNA (Fig. 18.18). Ambos os processos ocorrem no núcleo. Entretanto, antes que o RNA de **L1** seja reversamente transcrito, ele vai para o citoplasma onde é traduzido em polipeptídeos que aparentemente permanecem associados a ele quando ele retorna para o núcleo. O polipeptídeo codificado por ORF2 possui uma função de endonuclease que catalisa a clivagem de um filamento da dupla hélice de DNA em um sítio prospectivo de inserção em um cromossomo. A ponta 3' exposta deste filamento de DNA clivado serve então como um *primer* para a síntese de DNA usando o RNA **L1** como molde e a atividade da transcriptase reversa dada pelo polipeptídeo de ORF2. Deste modo, uma seqüência de DNA **L1** é sintetizada do ponto do cromossomo onde o polipeptídeo de ORF2 introduziu um corte unifilar. O DNA **L1** recém-sintetizado é subsequenteamente tornado bifilamentar pela síntese posterior de DNA, e o produto bifilamentar é então integrado covalentemente ao cromossomo, criando assim uma nova cópia do elemento **L1** no genoma. Às vezes, a região 5' do RNA de **L1** não é copiada em DNA. Quando isto acontece, a inserção **L1** resultante não terá as seqüências 5', isto é, será um elemento **L1** incompleto.

Apenas pequeno número dos elementos **L1** completos no genoma humano parece ser transposicionalmente ativo. Cópias transpostas destes elementos foram descobertas pela análise de doenças genéticas em um pequeno número de pessoas. Hoje em dia, 14 mutações de inserção de **L1** causadoras de doenças foram identificadas, incluindo mutações no gene do *fator VIII* (a causa de hemofilia) e no gene de *distrofina* (a causa de distrofia muscular). A transposição dos elementos **L1** parece, portanto, ser um evento raro. O genoma humano contém dois outros tipos de seqüências LINE, **L2** (315.000 cópias) e **L3** (37.000 cópias). Entretanto, nenhuma delas é transposicionalmente ativa.

Os **curtos elementos nucleares intercalares**, ou **SINES**, são a segunda classe mais abundante de elementos

de transposição no genoma humano. Tais elementos são tipicamente menores que 400 pares de bases de tamanho e não codificam proteínas. Como todos os transposons, eles têm uma seqüência de pares de bases A:T em uma ponta. Os **SINE** transpõem-se por um processo que envolve transcrição reversa de um RNA que foi transcrito de um promotor interno. Embora os detalhes do processo de transposição não sejam bem compreendidos, parece que a transcriptase reversa necessária para a síntese de DNA a partir do RNA de **SINE** é fornecida por um elemento tipo **LINE**. Assim, **SINE** dependem de **LINE** para se multiplicar e inserir no genoma. Neste sentido, elas podem ser consideradas como retroposons que são parasitas dos retroposons funcionalmente autônomos e autênticos. O genoma humano contém três famílias de **SINE**, a *Alu*, *MIR* e os elementos *Ther2/MIR3*. Entretanto, apenas os elementos *Alu*, denominados por uma enzima que reconhece uma seqüência específica de nucleotídeos dentro deles, são transposicionalmente ativos.

O genoma humano possui mais de 400.000 seqüências que são derivadas de elementos tipo retrovírus. A maioria destas seqüências é de LTR únicas, como as seqüências delta encontradas no genoma de leveduras. Embora já tenham sido identificadas mais de 100 famílias diferentes de elementos tipo retrovírus no DNA humano, apenas algumas parecem ter sido transposicionalmente ativas na história evolutiva recente. Como os **LINE** e **SINE** inativos, quase todas as seqüências tipo retrovírus humanas são fósseis genéticos abandonados de uma época em que tinham transposição ativa.

Transposons que são funcionalmente como *Ac/Ds* em milho, isto é, elementos de DNA que são transpostos por um mecanismo de cortar e colar catalisado por uma transposase codificada por elemento, são um pequeno componente do genoma humano. Foram identificados dois elementos que estão distantemente relacionados com os elementos *Ac/Ds* bem como alguns elementos na superfamília *Tc1/mariner*. Todas as evidências disponíveis indicam que estes tipos de transposons foram transposicionalmente inativos por muitos milhões de anos.

#### PONTOS IMPORTANTES

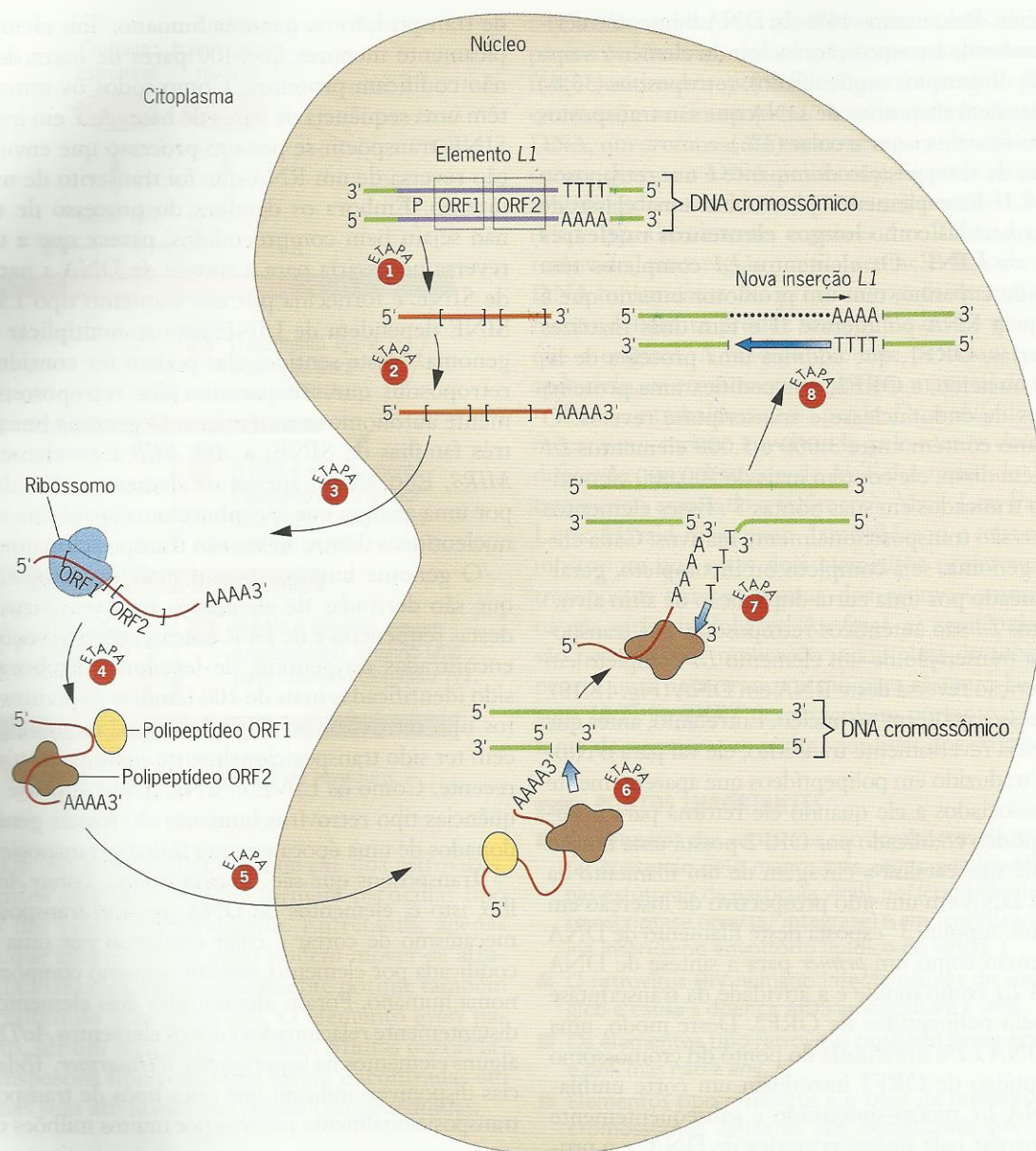
- O genoma humano contém quatro tipos básicos de elementos de transposição: **LINE**, **SINE**, elementos tipo retrovírus e transposons de recortar e colar.
- **LINE L1** e **SINE Alu** são transposicionalmente ativos; outros transposons humanos parecem ser inativos.

## O SIGNIFICADO GENÉTICO E EVOLUTIVO DE ELEMENTOS DE TRANSPOSIÇÃO

Os elementos de transposição moldaram genomas durante o curso da evolução.

Os elementos de transposição são muito dispersos e provavelmente são componentes antigos do genoma. Em algu-





**Fig. 18.18 ■** Mecanismo hipotético para a transposição de elementos L1 no genoma humano. O elemento L1 de aproximadamente 6 kb contém duas matrizes de leitura abertas, ORF1 e ORF2, transcritas de um promotor comum (P). O polipeptídeo codificado por ORF1 permanece associado ao RNA L1 e pode ser responsável por retornar o RNA para o núcleo. O polipeptídeo codificado por ORF2 tem pelo menos duas funções catalíticas. Primeira, é capaz de cortar os filamentos de DNA; logo, é uma endonuclease. Segunda, é capaz de sintetizar DNA a partir de um molde de RNA; logo, é uma transcriptase reversa. O tamanho da nova inserção L1 irá depender do quanto a transcriptase reversa viajar ao longo do molde de RNA L1. Se não conseguir atingir a ponta 5', a inserção será incompleta. Inserções incompletas geralmente têm promotores não funcionais e portanto não podem produzir RNA L1 para futuras transposições.

ETAPA 1. Um elemento L1 completo inserido em um cromossomo é transcrito em RNA L1.

ETAPA 2. O RNA L1 é poliadenilado no núcleo.

ETAPA 3. O RNA L1 poliadenilado move-se para o citoplasma.

ETAPA 4. O RNA L1 é traduzido em dois polipeptídeos correspondentes a cada uma das ORF. Esses polipeptídeos permanecem associados ao RNA L1.

ETAPA 5. O RNA L1 e seus polipeptídeos associados movem-se para o núcleo.

ETAPA 6. O polipeptídeo de ORF2 corta um filamento de uma molécula de DNA cromossômico, e a ponta 3' da cauda poli(A) no RNA L1 é justaposta ao lado 5' do DNA cortado.

ETAPA 7. O polipeptídeo de ORF2 exercita sua função de transcriptase reversa para sintetizar um único filamento de DNA usando o RNA L1 como molde. A ponta 3' do DNA cromossômico cortado serve como *primer* para esta síntese de DNA.

ETAPA 8. O único filamento recém-sintetizado de DNA muda para um local entre os dois sítios do DNA cromossômico cortado. Simultaneamente, o RNA L1 é eliminado, e o outro filamento do DNA cromossômico é cortado para permitir a síntese de um segundo filamento de DNA (linha pontilhada), complementar à sequência L1, no sentido indicado pela seta fina. Todos os cortes são reparados para ligar o elemento L1 recém-inserido ao DNA cromossômico.



mas espécies, eles constituem uma fração apreciável do DNA total. O DNA móvel é portanto um componente importante do genoma. Ele também contribui significativamente para a taxa total de mutação. Em *Drosophila*, por exemplo, talvez metade das mutações que ocorrem espontaneamente são causadas por inserções de elementos de transposição. Esta capacidade de causar danos mutacionais levanta questões sobre a condição evolutiva dos elementos de transposição. Eles têm alguma função útil ou são apenas parasitas genéticos, causando mutações ao perambular pelo genoma? De onde vieram os elementos de transposição? Que mecanismos evoluíram para controlar e limitar seu movimento?

### Transposons e Organização do Genoma

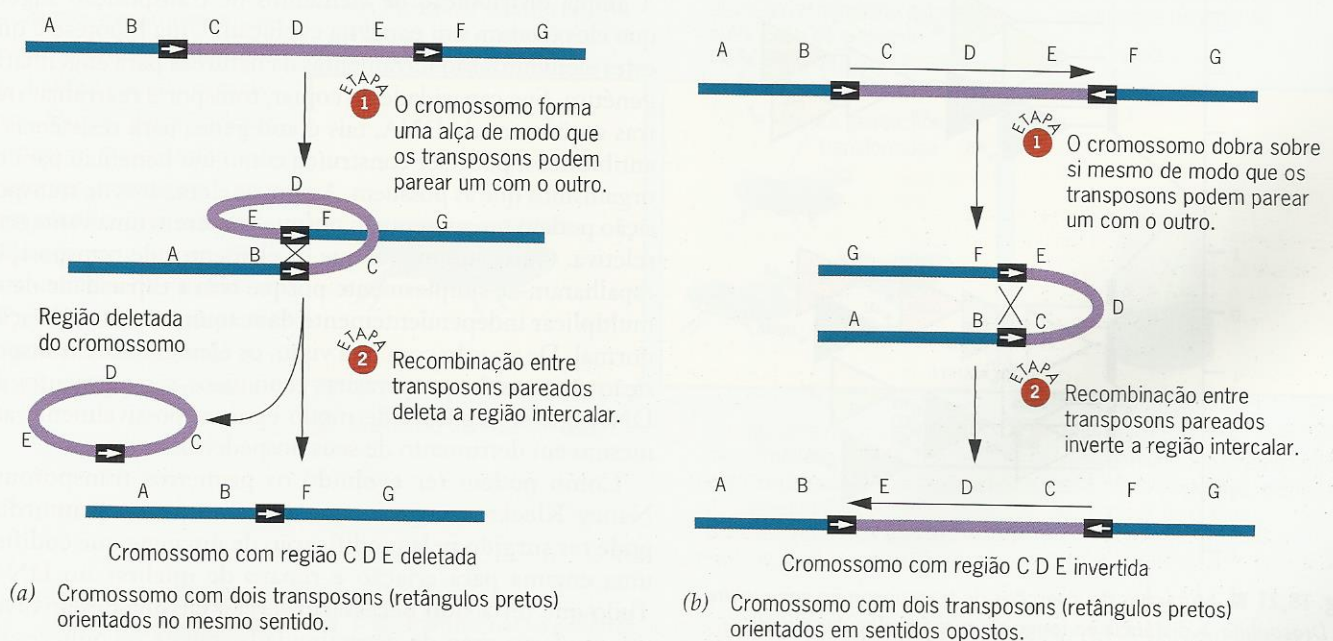
Algumas regiões cromossômicas são especialmente ricas em seqüências transposon. No milho, os elementos de transposição estão concentrados no DNA entre genes; coletivamente, esses elementos correspondem a mais da metade do DNA no genoma de milho. Em *Drosophila*, os transposons são concentrados na heterocromatina centromérica e na heterocromatina adjacente à eucromatina de cada braço cromossômico. Entretanto, muitos desses transposons mutaram a ponto de não poderem ser mobilizados. Geneticamente, eles são equivalentes a "mortos". A heterocromatina parece, portanto, ser um tipo de cemitério cheio de elementos de transposição degenerados.

Elementos de transposição também são encontrados na eucromatina, onde eles são dispersos em muitos locais diferentes. Em *Drosophila*, aproximadamente 90 famílias diferentes de elementos de transposição já foram identificadas.

O número de membros em cada família varia de alguns a uma ou duas centenas, mas a maioria das famílias parece ter entre 20 e 80 membros.

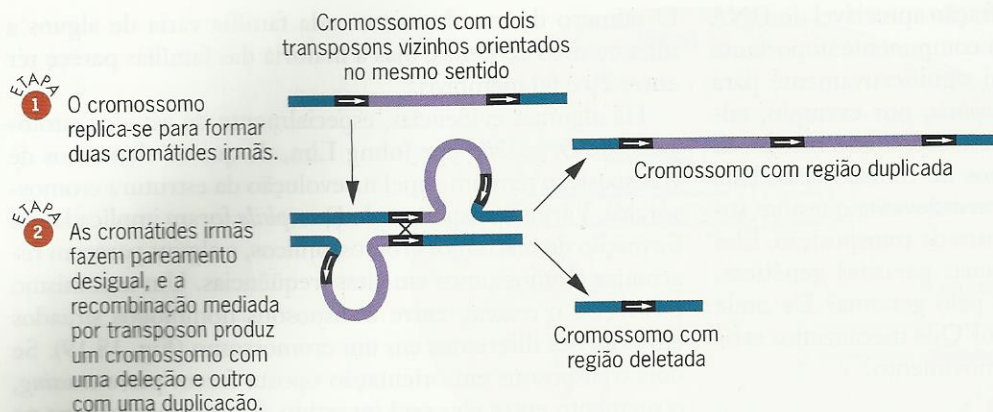
Há algumas evidências, especialmente de estudos citológicos de *Drosophila* por John Lim, de que os elementos de transposição têm um papel na evolução da estrutura cromossômica. Vários transposons de *Drosophila* foram implicados na formação de rearranjos cromossômicos, e alguns parecem rearranjar cromossomos em altas frequências. Um mecanismo possível é o *crossing* entre transposons homólogos situados em posições diferentes em um cromossomo (Fig. 18.19). Se dois transposons em orientação oposta fazem par e *crossing*, o segmento entre eles será invertido. Se dois transposons na mesma orientação fazem par e *crossing*, o segmento entre eles será deletado. Tais eventos são exemplos de **trocas intracromossômicas ectópicas**, isto é, trocas entre seqüências em sítios diferentes dentro de um único cromossomo. **Trocas intercromossômicas ectópicas** também são possíveis. Nestes tipos de eventos, seqüências em dois cromossomos diferentes (tanto homólogos quanto não homólogos) fazem par e *crossing*, gerando novos produtos. A Fig. 18.20 dá um exemplo de uma troca ectópica entre elementos de transposição em duas cromátides irmãs. Observe que um produto desta troca não tem o segmento entre os dois transposons, enquanto o outro contém uma duplicação deste segmento. Os elementos de transposição podem portanto mudar a estrutura cromossômica mediando trocas ectópicas.

Outro exemplo de troca ectópica é o evento de *crossing* que insere o plasmídeo F no cromossomo de *E. coli*. Tais eventos são mediados pelos elementos IS situados em ambas estas moléculas circulares.



**Fig. 18.19** ■ Rearranjos cromossômicos mediados por transposon. (a) Formação de uma deleção por recombinação intracromossômica entre dois transposons na mesma orientação. (b) Formação de uma inversão por recombinação intracromossômica envolvendo dois transposons em orientações opostas.





**Fig. 18.20** ■ Origem de uma duplicação gênica por *crossing* desigual mediado por transposon entre cromátides irmãs.

## Transposons e Mutação

Os elementos de transposição são responsáveis por mutações em uma grande variedade de organismos. A Fig. 18.21 mostra algumas das inserções de transposição que foram encontradas em alelos mutantes diferentes do gene *white* de *Drosophila*. Elas incluem vários tipos de elementos: *P*, ele-

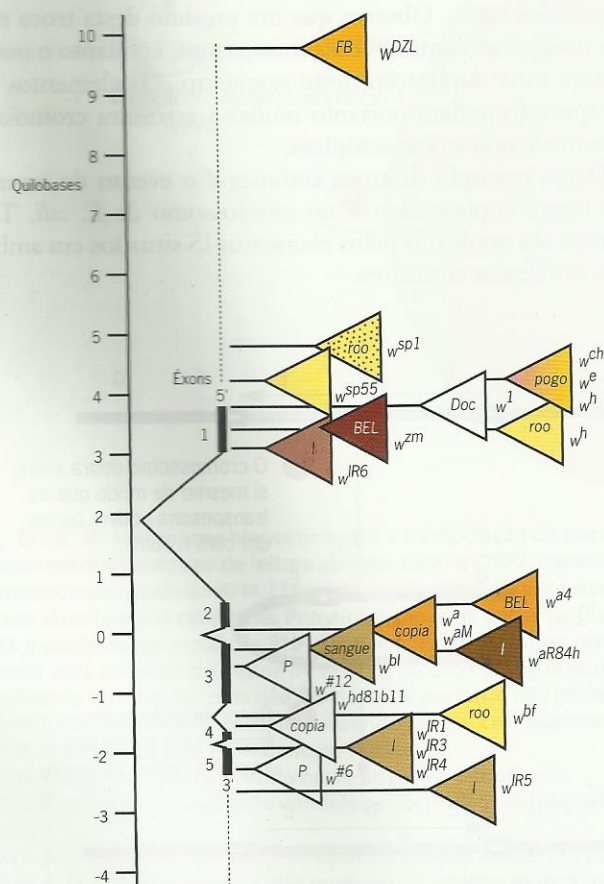
mentos tipo retrovírus e retroposons. Alguns desses elementos são inseridos em éxons, outros em íntrons e ainda outros em DNA regulador antecedente ao gene. O primeiro alelo mutante de *white*,  $w^1$ , descoberto por T. H. Morgan, resultou de uma inserção de transposon.

Embora inserções de transposons sejam comuns em estoques de organismos mutantes, a ocorrência de novas mutações de inserção é um evento raro. Isto sugere que o movimento de muitas famílias de transposon é regulado. Quando essa regulação é perturbada, pode ocorrer um surto de transposição, causando muitas mutações simultaneamente. Isto é aparentemente o que ocorre quando elementos *P* são mobilizados em híbridos disgênicos de *Drosophila*.

## Aspectos Evolutivos Envolvendo Elementos de Transposição

A ampla distribuição de elementos de transposição sugere que eles tiveram um papel na evolução. Uma hipótese é que estes elementos são ferramentas da natureza para engenharia genética. Sua capacidade de copiar, transpor e rearranjar outras seqüências de DNA, tais como genes para resistência a antibióticos, pode ser construída como um benefício para os organismos que as possuem. Assim, os elementos de transposição podem ter aumentado porque conferem uma vantagem seletiva. Outra hipótese é que os elementos de transposição espalharam-se simplesmente porque têm a capacidade de se multiplicar independentemente da maquinaria de replicação normal. De acordo com esta visão, os elementos de transposição são mais do que parasitas genômicos, são segmentos de DNA que se replicam de modo egoísta, possivelmente até mesmo em detrimento de seus hospedeiros.

Como podem ter evoluído os primeiros transposons? Nancy Kleckner sugeriu que um transposon primordial pode ter surgido pela modificação de um gene que codifica uma enzima para criação e reparo de quebras no DNA. Tudo que teria sido necessário seria a enzima desenvolver um modesto grau de especificidade, talvez reconhecendo uma determinada seqüência de DNA com seis ou oito pares de nucleotídeos. Tal seqüência pode ocorrer por acaso em orientação invertida em um dos lados do gene, criando



**Fig. 18.21** ■ Mutações de inserções de transposon no gene *white* de *Drosophila*. A distância ao longo do mapa molecular é dada em pares de quilobases, com a coordenada zero arbitrariamente posicionada no segundo íntron do gene no sítio de inserção *copia* na mutação  $w^a$ . Cada triângulo representa uma inserção de transposon diferente no gene *white*. As cores indicam mutantes de cor de olhos.

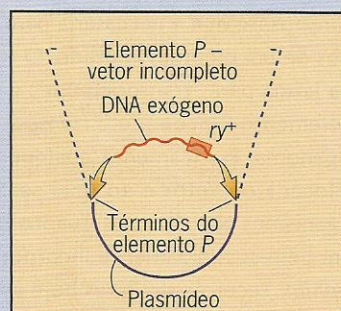




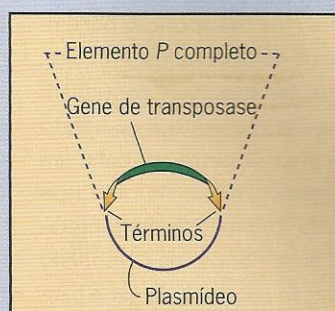
## UM MARCO NA GENÉTICA: Transformação de *Drosophila* com Elementos

Nos anos 1940, Oswald Avery e colaboradores descobriram que bactérias podiam ser geneticamente alteradas quando tratadas com DNA isolado. Neste processo, um fragmento de DNA entra na célula e é fisicamente recombinado com o cromossomo. A célula recombinante pode então ser cultivada para produzir uma linhagem de organismos geneticamente transformados. A descoberta de Avery criou a esperança de que algum dia seria possível alterar os genomas de organismos eucarióticos inserindo-se neles fragmentos de DNA específico. Quando aplicado a seres humanos, tal procedimento pode fornecer um modo de corrigir doenças genéticas.

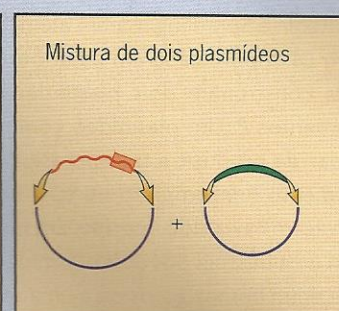
Por quase 40 anos, a produção experimental de transformantes genéticos foi limitada a microrganismos. Muitos pesquisadores tentaram transformar eucariotes superiores, mas sem sucesso. Esta cadeia de falhas foi rompida em 1982, quando Gerald Rubin e Allan Spradling produziram a primeira *Drosophila*<sup>1</sup> geneticamente transformada.



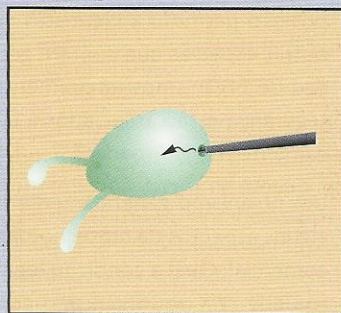
**ETAPA 1** DNA exógeno é inserido em um elemento P incompleto em um plasmídeo. O inserto também contém um gene de cor de olho ( $ry^+$ ) como marcador. Em moscas, este gene produz olhos vermelhos.



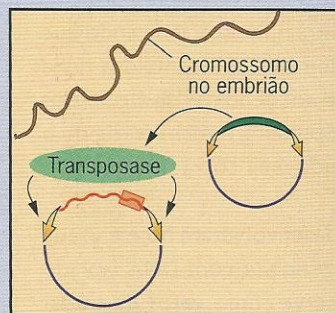
**ETAPA 2** Um elemento P completo é inserido em um plasmídeo diferente.



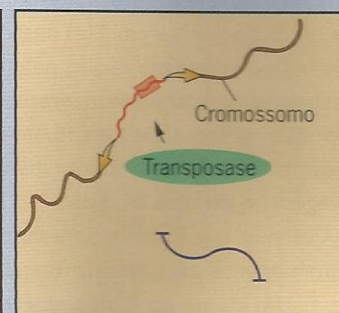
**ETAPA 3** Os dois plasmídeos são misturados na solução.



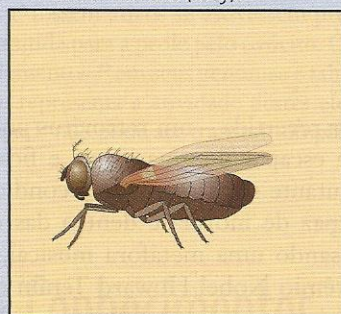
**ETAPA 4** A mistura de plasmídeos é microinjetada em um embrião de *Drosophila* mutante  $ry^-$ . As moscas com a mutação  $ry^-$  têm olhos marrons (*rosy*).



**ETAPA 5** Na linhagem germinativa do embrião, a transposase do elemento P completo catalisa a remoção do elemento P incompleto de seu plasmídeo.



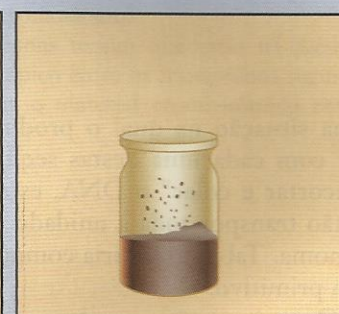
**ETAPA 6** O elemento P excisado é inserido em um cromossomo na linhagem germinativa do embrião.



**ETAPA 7** O embrião torna-se um adulto.



**ETAPA 8** Quando uma mosca adulta se reproduz, a prole é examinada quanto à incorporação do elemento P incompleto com seu gene marcador  $ry^+$ .



**ETAPA 9** Prole geneticamente transformada, reconhecida por seus olhos vermelhos, é criada em frascos de cultura.

<sup>1</sup>Rubin, G. M., e A. C. Spradling, 1982. Genetic transformation of *Drosophila* with transposable element vectors. *Science* 218: 348-353.

**Fig. 1** ■ Transformação genética de *Drosophila* usando vetores do elemento P. DNA exógeno inserido entre as pontas do elemento P é integrado ao genoma pela ação de uma transposase codificada pelo elemento P completo. Moscas com este DNA em seus genomas podem ser reproduzidas em culturas de laboratório.





## UM MARCO NA GENÉTICA (continuação)

Rubin e Spradling usaram elementos *P* de transposição para inserir DNA purificado em embriões vivos de *Drosophila*. Primeiro, eles construíram dois plasmídeos bacterianos que continham elementos *P*. Um plasmídeo continha um elemento *P* completo capaz de produzir a transposase *P in vivo*. O outro continha um elemento *P* incompleto no qual um gene para cor de olho tipo selvagem havia sido inserido. Em seguida, Rubin e Spradling injetaram uma mistura dos dois plasmídeos em embriões de *Drosophila* que eram homozigotos para uma mutação recessiva do gene para cor de olho. Eles esperavam que a transposase produzida pelo elemento *P* completo fosse catalisar o elemento *P* incompleto para pular de seu plasmídeo para os cromossomos dos animais injetados, levando com ele o gene para cor de olho. Quando esses animais fossem adultos, Rubin e Spradling os cruzariam com moscas homozigotas para a mutação de cor de olho e veriam se a prole tinha olhos tipo selvagem. Eles descobriram muitos, indicando que o gene tipo selvagem levado pelo elemento *P* incompleto tinha sido bem-sucedidamente incorporado aos genomas de alguns dos embriões injetados. De fato, Rubin e Spradling haviam corrigido o fenótipo cor de olho mutante inserindo uma cópia do gene tipo selvagem no genoma da mosca.

A técnica que Rubin e Spradling desenvolveram hoje é usada rotineiramente para transformar *Drosophila* com DNA isolado (Fig. 1). Um elemento *P* incompleto serve como **vetor de transformação** e um elemento *P* completo serve como a fonte da

transposase que é necessária para inserir o vetor nos cromossomos de um embrião injetado. O termo *vetor* origina-se da palavra latina para “transportador”; é usado neste contexto porque o elemento *P* incompleto *transporta* um fragmento de DNA para o genoma. Praticamente qualquer sequência de DNA pode ser colocada no vetor e ao final ser inserida no animal. Genes de organismos tão diversos quanto bactérias e seres humanos foram bem-sucedidamente incorporados a cromossomos de *Drosophila*. Técnicas similares hoje estão sendo desenvolvidas para a transformação genética de outros organismos, incluindo nossa própria espécie.

### QUESTÕES PARA DISCUSSÕES

1. A transformação genética com vetores de elementos de transposição permitiu a cientistas criar linhagens de *Drosophila* e outras espécies que contêm genes “exógenos” – genes derivados de organismos inteiramente diferentes. Que aspectos de segurança estão associados à criação destas linhagens geneticamente transformadas?
2. Um organismo geneticamente transformado pode ser útil em alguns aspectos da biotecnologia e pode portanto ser de valor comercial. Qual a condição legal de tais organismos? Uma pessoa que cria um organismo geneticamente modificado pode obter uma patente de tal organismo?

uma situação na qual o produto gênico poderia interagir com cada uma destas sequências flanqueadoras. Por “recortar e colar” o DNA, esta enzima modificada pode então transpor toda a unidade para uma nova posição no genoma. Tal unidade iria comportar-se como um transposon primitivo.

Outras questões envolvem a correlação evolutiva entre elementos tipo retrovírus, tais como *Ty1* em leveduras, e retrovírus desenvolvidos. Coletivamente, estas entidades foram chamadas de **retroelementos**. Alan Kingsman e Susan Kingsman propuseram que os retrovírus desenvolveram-se a partir de retrotransposons mais simples pela adição de um gene (chamado *env*) que sintetiza uma proteína de membrana. Com esta adição, o retroelemento poderia produzir uma partícula capaz de escapar de uma célula e entrar em outra. Tal partícula seria infecciosa e portanto daria ao retroelemento a oportunidade de se transpor entre os genomas bem como dentro deles. Logicamente, a situação poderia

ser revertida – um retrovírus poderia perder sua capacidade de escapar de uma célula e ficar capturado nela. Tal vírus mutante seria reduzido à condição de um retrotransposon, capaz de mover-se dentro das células mas não entre elas. Usando uma metáfora musical, o virologista ganhador do Prêmio Nobel Howard Temin uma vez descreveu estes cenários contrastantes como escalas evolutivas ascendentes e descendentes. Os retrotransposons podem subir ao nível dos retrovírus, e os retrovírus podem cair ao nível dos retrotransposons.

### PONTOS IMPORTANTES

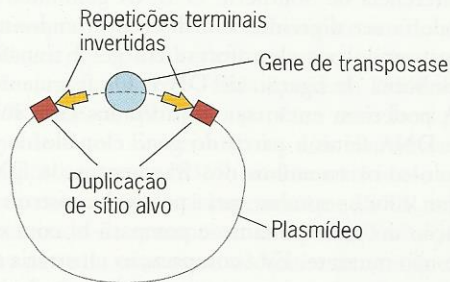
- Transposons são encontrados nos genomas de muitos tipos de organismos.
- Transposons são uma causa importante de mutações e rearranjos cromossômicos.
- Alguns transposons conferem uma vantagem seletiva aos organismos que os possuem; outros são parasitas genéticos.



## Exercícios Básicos Ilustram a análise genética básica.

1. Desenhe um elemento *IS* bacteriano inserido em um plasmídeo circular. Indique as posições (a) do gene de transposase, (b) as repetições terminais invertidas e (c) o sítio alvo da duplicação.

**Resposta:**



2. Que fator deve estar presente em milho para mobilizar um elemento *Ds* inserido em um braço cromossômico?

**Resposta:** Um elemento *Ds* é mobilizado quando a transposase codificada por um elemento *Ac* atua nele. Um elemento *Ac* deve portanto estar presente no genoma de milho.

3. Um geneticista tem duas linhagens de *Drosophila*. Uma, um estoque de longa duração em laboratório com olhos brancos, não tem elementos *P*; a outra, recentemente derivada de moscas tipo selvagem coletada em um mercado de frutas, tem elementos *P* em seu genoma. Quais dos seguintes cruzamentos esperar-se-ia que produzissem prole híbrida disgênica: (a) fêmeas de olhos brancos × machos tipo selvagem, (b) machos de olhos brancos × fêmeas tipo selvagem, (c) fêmeas brancas × machos brancos, (d) fêmeas tipo selvagem × machos tipo selvagem?

**Resposta:** (a) Fêmeas de olhos brancos × machos tipo selvagem. As fêmeas de olhos brancos não têm os elementos *P* em seus genomas e têm o citotipo M, que é a condição celular que permitiria que os elementos *P* fossem ativos se estivessem presentes. Esta condição é transmitida para a prole pelo ovócito. Os machos tipo selvagem têm elementos *P* em seus genomas e também têm o citotipo P, o que representa a atividade destes elementos. Entretanto, o citotipo P não pode ser transmitido pelo

espermatozóide. Assim, quando os machos tipo selvagem são cruzados com fêmeas de olhos brancos, a prole herda elementos *P* de seus pais e o citotipo M de suas mães. Esta combinação de fatores permite que os elementos *P* herdados paternamente se tornem ativos nos tecidos da linhagem germinativa da prole e ocorra disgenesia híbrida.

4. Quais as similaridades e diferenças entre retrovírus, elementos tipo retrovírus e transposons?

**Resposta:** Todos os três tipos de retroelementos usam transcrição reversa para inserir cópias de DNA de seu RNA em novos sítios do genoma da célula. Além disso, a enzima (transcriptase reversa) que catalisa a transcrição reversa é codificada por cada tipo de elemento. Para retrovírus e elementos tipo retrovírus, a transcrição reversa do RNA ocorre no citoplasma, enquanto para retrotransposons ela ocorre no núcleo. Retrovírus e elementos tipo retrovírus codificam outra proteína que funciona na montagem do vírus ou partículas tipo vírus no citoplasma. Os retrotransposons codificam uma proteína diferente que parece ligar-se ao RNA retroposon e levá-lo para o núcleo. O RNA retroviral é embalado em partículas virais, que podem sair da célula. Esta capacidade de sair requer uma proteína codificada pelo gene *env* no genoma viral. Como nem os elementos tipo retrovírus nem os retrotransposons levam um gene *env*, seu RNA não pode ser embalado para sair da célula. Retrovírus são infecciosos; elementos tipo retrovírus e retrotransposons não.

5. Que elemento de transposição é mais abundante no genoma humano?

**Resposta:** O LINE conhecido como *L1* é o transposon humano mais abundante. Ele corresponde a cerca de 17% de todo o DNA humano.

6. Como podem dois transposons na mesma família causar deleção de DNA entre eles em um cromossomo?

**Resposta:** Os dois transposons teriam que estar na mesma orientação. Pareamento ectópico entre os transposons seguido de recombinação removeria o material cromossômico entre eles. Veja a Fig. 18.16a.

## Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. Uma cópia do gene *white* do tipo selvagem ( $w^+$ ) de *Drosophila* foi inserida no meio de um elemento *P* incompleto contido dentro de um plasmídeo. O plasmídeo foi misturado com outro plasmídeo que contém um elemento *P* completo, e a mistura foi cuidadosamente injetada em embriões de *Drosophila* homozigotos para uma mutação nula ( $w^-$ ) do gene *white*. Os adultos que se desenvolveram desses embriões injetados tinham todos olhos brancos, mas, quando foram cruzados com moscas de olhos brancos, parte de sua prole tinha olhos vermelhos. Explique a origem desta prole de olhos vermelhos.

**Resposta:** O elemento *P* completo em um dos plasmídeos produziria a transposase *P*, a enzima que catalisa a transposição do elemento *P*, nas linhagens germinativas dos embriões injetados. O elemento *P* incompleto no outro plasmídeo seria um alvo para esta transposase. Se tal elemento *P* incompleto fosse mobilizado pela transposase para saltar de seu plasmídeo para os cromossomos do embrião injetado, a mosca que se desenvolveria deste embrião levaria uma cópia do gene *white* do tipo selvagem em sua linhagem germinativa. (O movimento do elemento *P* é limitado à linhagem germinativa. Portanto, o elemento *P*



incompleto não saltaria para os cromossomos das células somáticas, tais como as que por fim formam os olhos.) Tal mosca geneticamente transformada teria de fato o genótipo  $w^-/w^-$  na linhagem germinativa;  $P(w^+)$  ou  $w^-/Y$ ;  $P(w^+)$  indica o elemento  $P$  incompleto que contém o gene  $w^+$ . Este elemento poderia ser inserido em qualquer um dos cromossomos. Se a mosca transformada fosse cruzada com uma mosca de olhos brancos não injetada, parte de sua prole herdaria a inserção  $P(w^+)$ , a qual, como leva o gene *white* tipo selvagem, causaria o desenvolvimento de olhos vermelhos. A prole de olhos vermelhos é portanto o resultado de transformação genética de uma mosca mutante pelo gene  $w^+$  dentro do elemento  $P$  incompleto.

2. O elemento *Alu* é um dos SINE no genoma humano. Cada retroposon *Alu* tem cerca de 300 pares de bases de tamanho, tamanho insuficiente para codificar uma transcriptase reversa que possa catalisar a conversão de RNA *Alu* em DNA *Alu* durante o processo de retrotransposição. Apesar desta deficiência, os elementos *Alu* acumularam-se tanto que constituem 11% do DNA humano – mais de 1 milhão de cópias. Como esta incrível expansão de elementos *Alu* ocorreu durante a história evolutiva da linhagem humana sem uma transcriptase reversa codificada por *Alu*?

**Resposta:** Os elementos *Alu* podem ter “pedido emprestado” os serviços de uma transcriptase reversa codificada por um retroposon diferente tal como o elemento *L1*, que é grande o suficiente para codificar uma transcriptase reversa e pelo menos outro polipeptídeo. Se a transcriptase reversa codificada por *L1*, ou a transcriptase reversa codificada por algum outro retroelemento, talvez outro LINE, puder ligar-se ao RNA *Alu*, então é concebível que a transcriptase reversa possa usar o RNA *Alu* para sintetizar DNA *Alu*, que poderia subsequentemente integrar-se ao DNA cromossômico. A repetição deste processo durante o tempo evolutivo poderia explicar o acúmulo de tantas cópias do elemento *Alu* no genoma humano.

3. Que técnicas podem ser usadas para demonstrar que uma mutação em um ser humano com hemofilia é devida à inserção de um elemento *Alu* na sequência codificante do gene ligado ao X para o fator VIII, que é uma das proteínas necessárias para coagulação sanguínea eficiente em seres humanos?

**Resposta:** Um geneticista molecular teria vários modos para mostrar que o gene mutante para hemofilia é devido a uma inserção *Alu* na sequência codificante do gene. Uma técnica é a transferência de Southern. O DNA genômico do hemofílico poderia ser digerido com diferentes endonucleases de restrição, partido por eletroforese em gel e transferido para uma membrana de ligação de DNA. Os fragmentos ligados de DNA poderiam então ser hibridizados com sondas marcadas de DNA feitas a partir do gene clonado de *fator VIII*. Analisando-se os tamanhos dos fragmentos de DNA que se hibridizam com as sondas, seria possível construir um mapa de restrição do gene mutante e compará-lo com o mapa de um gene não mutante. Esta comparação mostraria a presença de uma inserção no gene mutante. Também poderia revelar a identidade da sequência inserida. (Elementos *Alu* são clivados por uma determinada enzima de restrição, *Alu I*, que pode ser uma das enzimas usadas na análise.) Uma técnica mais simples é amplificar partes da sequência codificante do gene do *fator VIII* usando-se a reação em cadeia da polimerase (PCR). Pares de *primers* posicionados apropriadamente ao longo da sequência codificante poderiam ser usados em uma série de reações de amplificação, cada uma das quais seria conectada com o DNA molde de um hemofílico. Cada par de *primers* deveria amplificar um segmento do gene do *fator VIII*. Os tamanhos dos produtos de PCR poderiam então ser determinados por eletroforese em gel. Uma inserção *Alu* em um segmento particular do gene aumentaria o tamanho deste segmento em cerca de 300 pares de bases. Esta suposta inserção *Alu* poderia ser identificada definitivamente sequenciando-se o DNA do produto de PCR maior que o normal.

## Questões e Problemas Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 18.1. Qual dos seguintes pares de sequências de DNA se qualifica como repetições terminais de um elemento IS bacteriano:  
(a) 5'-GAATCCGCA-3' e 5'-ACGCCTAAG-3',  
(b) 5'-GAATCCGCA-3' e 5'-CTTAGGCGT-3',  
(c) 5'-GAATCCGCA-3' e 5'-GAATCCGCA-3',  
(d) 5'-GAATCCGCA-3' e 5'-TGCGGATTC-3'. Explique.

- 18.2. Qual dos seguintes pares de sequências de DNA se qualifica como duplicações de sítio alvo no ponto de uma inserção IS50?  
(a) 5'-AATTCGCGT-3' e 5'-AATTCGCGT-3',  
(b) 5'-AATTCGCGT-3' e 5'-TGCGCTTAA-3',  
(c) 5'-AATTCGCGT-3' e 5'-TTAAGCGCA-3',  
(d) 5'-AATTCGCGT-3' e 5'-ACGCGAATT-3'. Explique.

- 18.3. Uma linhagem de *E. coli* é resistente ao antibiótico estreptomicina e outra linhagem é resistente ao antibiótico ampicilina. As duas linhagens foram cultivadas juntas e então plaqueadas em meio seletivo contendo estreptomicina e

ampicilina. Surgiram várias colônias, indicando que as células adquiriram resistência a ambos os antibióticos. Sugira um mecanismo para explicar a aquisição da dupla resistência.

- 18.4. O que distingue elementos IS e Tn3 em bactérias?

- 18.5. A ordem circular dos genes no cromossomo de *E. coli* é  $*A B C D E F G H *$ , onde  $*$  indica que as pontas dos cromossomos estão ligadas uma à outra. Duas cópias de um elemento IS estão situadas neste cromossomo, uma entre os genes *C* e *D* e a outra entre os genes *D* e *E*. Uma única cópia deste elemento também está presente no plasmídeo F. Duas linhagens Hfr foram obtidas selecionando-se a integração do plasmídeo F ao cromossomo. Durante a conjugação, uma linhagem transfere os genes cromossômicos na ordem *D E F G H A B C*, enquanto a outra os transfere na ordem *D C B A H G F E*. Explique a origem destas duas linhagens Hfr. Por que elas transferem genes em ordens diferentes? A



ordem de transferência revela algo sobre a orientação dos elementos IS no cromossomo de *E. coli*?

- 18.6. Ao acaso, um elemento IS1 foi inserido perto de um elemento IS2 no cromossomo de *E. coli*. O gene entre eles, *sug*<sup>+</sup>, confere a capacidade de metabolizar alguns açúcares. A unidade IS1 *sug*<sup>+</sup> IS2 irá comportar-se como um transposon composto? Explique.
- 18.7. O transposon composto Tn5 consiste em dois elementos IS50, um de cada lado de um grupo de três genes para resistência a antibióticos. Toda a unidade IS50L *kan*<sup>r</sup> *ble*<sup>r</sup> *str*<sup>r</sup> IS50R pode ser transposta para um novo local no cromossomo de *E. coli*. Entretanto, dos dois elementos IS50 neste transposon, apenas IS50R produz a atividade catalítica de transposase ativa. Você esperaria que IS50R seja capaz de ser excisado do transposon composto Tn5 e inserido em outra parte do cromossomo? Você esperaria que IS50L seja capaz de fazer isto?
- 18.8. Um pesquisador encontrou um novo elemento Tn5 com a estrutura IS50L *str*<sup>r</sup> *ble*<sup>r</sup> *kan*<sup>r</sup> IS50L. Qual a origem mais provável deste elemento?
- 18.9. Um elemento Tn3 com uma mutação de mudança de matriz de leitura no início do gene *tnpA* seria capaz de formar um co-integrado? Um elemento Tn3 com uma mutação de mudança de matriz de leitura no início do gene *tnpR* seria capaz de formar um co-integrado?
- 18.10. Que enzimas são necessárias para transposição replicativa de Tn3? Quais são suas funções respectivas?
- 18.11. Qual o significado médico dos transposons bacterianos?
- 18.12. Descreva a estrutura do transposon *Ac* no milho. De que modos os transposons *Ds* diferem estruturalmente e funcionalmente do transposon *Ac*?
- 18.13. Na condição homozigota, uma mutação de deleção do locus *c*, *c*<sup>n</sup>, produz grãos incolores (brancos) em milho. O alelo dominante tipo selvagem, *C*, faz com que os grãos fiquem púrpura. Uma mutação recessiva recém-identificada do locus *c*, *c*<sup>m</sup>, tem o mesmo fenótipo que a mutação de deleção (grãos brancos), mas, quando são cruzadas as plantas *c*<sup>m</sup> *c*<sup>m</sup> e plantas *c*<sup>n</sup> *c*<sup>n</sup>, elas produzem grãos brancos com listras púrpura. Se soubermos que as plantas *c*<sup>n</sup> *c*<sup>n</sup> possuem elementos *Ac*, qual a explicação mais provável para a mutação *c*<sup>m</sup>?
- 18.14. Em milho, o gene *O2*, situado no cromossomo 7, controla a textura do endosperma, e o gene *C*, situado no cromossomo 9, controla sua cor. O gene no cromossomo 7 tem dois alelos, um recessivo, *o2*, que faz com que o endosperma fique macio, e um dominante, *O2*, que faz com que ele seja duro. O gene no cromossomo 9 também tem dois alelos, um recessivo, *c*, que permite que o endosperma seja colorido, e um dominante, *C*<sup>d</sup>, que inibe a coloração. Em uma linhagem homozigota *C*<sup>d</sup>, um elemento *Ds* é inserido no cromossomo 9 entre o gene *C* e o centrômero. Este elemento pode ser ativado introduzindo-se um elemento *Ac* por cruzamentos apropriados. A ativação de *Ds* faz com que o alelo *C*<sup>d</sup> seja perdido por quebra cromossômica. Nos grãos *C*<sup>d</sup>/*c*, tal perda produz manchas de tecido colorido em um fundo de outro modo incolor (veja Figs. 18.8 e 18.9). Um geneticista cruza uma linhagem com o genótipo *o2/o2*; *C*<sup>d</sup> *Ds*/*C*<sup>d</sup> *Ds* com uma linhagem com o genótipo *O2/o2*; *c/c*. Esta última linhagem leva um elemento *Ac* em algum lugar no genoma. Entre a prole, apenas aqueles com endosperma duro apresentam manchas de tecido colorido. O que isto lhe diz sobre a localização do elemento *Ac* na linhagem *O2/o2*; *c/c*?
- 18.15. Em milho, o alelo recessivo *bz* (*bronze*) produz uma cor mais clara na aleurona do que o alelo dominante, *Bz*. Uma planta homozigota *bz/bz* foi fertilizada com o pólen de uma planta homozigota *Bz/Bz*. Os sabugos resultantes continham grãos que eram uniformemente escuros exceto alguns em que ocorriam pontos claros. Sugira uma explicação.
- 18.16. O locus *singed* ligado ao X é um dos vários em *Drosophila* que controlam a formação de cerdas na cutícula adulta. Machos que são hemizigotos para um mutante do alelo *singed* têm cerdas dobradas torcidas que em geral são de tamanho reduzido. Várias mutações de inserção de elemento *P* do locus *singed* já foram caracterizadas, e se demonstrou que algumas reverterem para o alelo tipo selvagem por excisão do elemento inserido. Que condições devem estar presentes para permitir que ocorram tais reversões?
- 18.17. Híbridos disgênicos em *Drosophila* têm elevadas taxas de mutação como resultado de transposição do elemento *P*. Como você aproveitaria esta situação para obter mutações por inserção de elemento *P* no cromossomo X?
- 18.18. Se o DNA de uma mutação de inserção de elemento *P* do gene *white* de *Drosophila* e o DNA de um gene *white* tipo selvagem fossem purificados, desnaturados, misturados um com o outro, renaturados e então vistos ao microscópio eletrônico, como pareceriam as moléculas híbridas de DNA?
- 18.19. Quando elementos *P* completos são injetados em embriões de uma linhagem M, eles são transpostos para cromossomos da linhagem germinativa, e a prole gerada por estes embriões pode ser usada para estabelecer novas linhagens P. Entretanto, quando elementos *P* completos são injetados em embriões de insetos que não têm estes elementos, tais como mosquitos, eles não são transpostos para os cromossomos da linhagem germinativa. O que esta falha na inserção em cromossomos de outros insetos indica sobre a natureza da transposição de elementos *P*?
- 18.20. (a) O que são elementos tipo retrovírus? (b) Cite exemplos de elementos tipo retrovírus em leveduras e *Drosophila*. (c) Descreva a forma de transposição dos elementos tipo retrovírus? (d) Após um elemento tipo retrovírus ter sido inserido em um cromossomo, espera-se que ele seja excisado?
- 18.21. Às vezes, cópias únicas de LTR de um retrotransposon chamado *gypsy* são encontradas em cromossomos de *Drosophila*. Como se originaram essas cópias únicas de LTR?



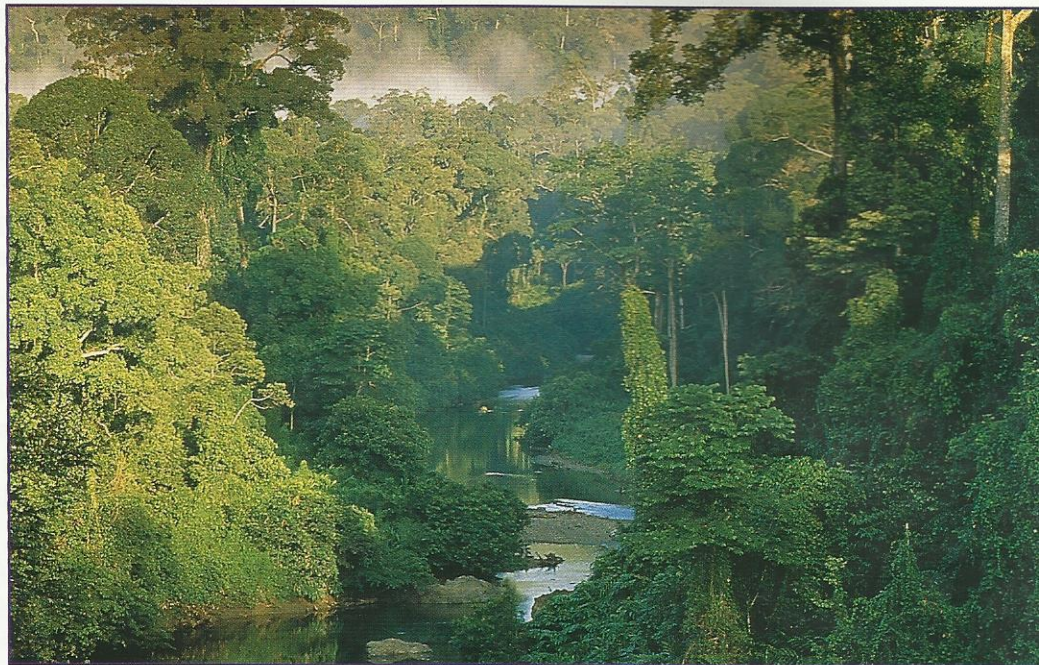
- 18.22. Você esperaria que genes em um retrotransposon possuam íntrons? Explique.
- 18.23. Sugira um método para determinar se um retroposon *TART* está situado nos telômeros de cada um dos cromossomos no genoma de *Drosophila*.
- 18.24. Foi proposto que os elementos de transposição *hobo* em *Drosophila* mediem recombinação intracromossômica, isto é, dois elementos *hobo* no mesmo cromossomo formem um par e se recombinem. O que tal evento de recombinação produziria se os elementos *hobo* fossem orientados no mesmo sentido no cromossomo? E se fossem orientados em sentidos opostos?
- 18.25. Que evidência sugere que alguns elementos de transposição não são simples parasitas genéticos?
- 18.26. Aproximadamente metade de todas as mutações espontâneas em *Drosophila* são causadas por inserções de elementos de transposição. Em seres humanos, entretanto, evidências acumuladas sugerem que a maioria das mutações espontâneas não são causadas por inserções de transposons. Propõe uma hipótese para explicar a diferença.
- 18.27. Z. Ivics, Z. Izsvák e P. B. Hackett “ressuscitaram” um elemento não móvel da família *Tc1/mariner* de elementos de transposição isolados do DNA de salmão. Estes pesquisadores alteraram 12 códons dentro da sequência codificante do gene de transposase do elemento de salmão para restaurar a função catalítica de sua transposase. O elemento alterado, chamado *Sleeping Beauty*, está sendo testado como um agente para a transformação genética de vertebrados tais como camundongos e peixe-zebra (e possivelmente seres humanos). Suponha que você tem um plasmídeo bacteriano contendo o gene para a proteína verde fluorescente (*gfp*) inserido entre as pontas do elemento *Sleeping Beauty*. Como você faria para obter camundongos ou peixe-zebra que expressem o gene *gfp*?
- 18.28. O genoma humano contém cerca de 5.000 “pseudogenes processados”, que são derivados da inserção de cópias de DNA de moléculas de mRNA derivadas de muitos genes diferentes. Preveja a estrutura desses pseudogenes. Seria esperado que cada tipo de pseudogene processado tenha uma nova família de retrotransposons dentro do genoma humano? Pode-se esperar que o número de cópias de cada tipo de pseudogene processado aumente significativamente com o tempo evolutivo, como ocorreu com o número de cópias da família *Alu*? Explique suas respostas.



# A Genética de Mitocôndrias e Cloroplastos

## Tópicos

- A Genética Clássica das Organelas
- A Genética Molecular de Mitocôndrias
- DNA Mitocondrial e Doenças Humanas
- A Genética Molecular dos Cloroplastos
- Origem e Evolução de Mitocôndrias e Cloroplastos



Uma floresta tropical é uma acentuada manifestação de conversões de energia que ocorrem em cloroplastos e mitocôndrias.

## Mitocôndrias, Cloroplastos e a Engrenagem Energética Biológica

A Terra primitiva era um mundo muito diferente do que conhecemos hoje. O planeta era despojado de qualquer organismo vivo, e o ar consistia primariamente em metano, amônia e vapor de água. O planeta provavelmente era muito mais quente do que é hoje em dia, irradiando grandes quantidades de calor de seu cerne de lava, e ocorriam violentas erupções vulcânicas bem como intensas tempestades elétricas. Radiação ultravioleta letal banhava a superfície do planeta, energizando reações químicas no ar e acúmulos rasos de água. Nesse ambiente hostil começou a vida, provavelmente como um aglomerado de moléculas orgânicas que surgiam da fusão de compostos mais simples. Posteriormente, a organização celular desenvolveu-se e se diversifi-



cou, resultando finalmente em um grupo de organismos que podiam pensar, falar e conduzir pesquisas científicas. Motivados por uma intensa curiosidade, esses organismos tentaram reconstituir o longo curso da história e reconstruir os eventos sombrios de suas próprias origens.

Esta pesquisa revelou que os primeiros organismos verdadeiros foram células unitárias que viveram na água e se alimentaram de materiais orgânicos no ambiente. Tais organismos sustentam-se fermentando açúcares simples, tais como glicose, e gerando dióxido de carbono ( $\text{CO}_2$ ) como subproduto. O acúmulo de dióxido de carbono proporcionou um novo recurso ao planeta primitivo, e surgiram organismos que puderam explorá-lo. Esses organismos desenvolveram um mecanismo para sintetizar açúcares simples a partir de dióxido de carbono e água captando energia solar, o processo chamado fotossíntese. Os açúcares criados pela fotossíntese foram então fermentados para fornecer energia para a vida. À medida que a fotossíntese se desenvolveu e se espalhou, foi produzido um novo gás, o oxigênio ( $\text{O}_2$ ), que então permitiu a evolução de um novo grupo de organismos. Tais organismos foram capazes de usar o oxigênio produzido pela fotossíntese para extrair energia adicional de açúcares por meio de um processo químico chamado metabolismo oxidativo, ou aeróbico. O metabolismo aeróbico superou o processo anaeróbico mais simples de fermentação e resultou na evolução de organismos mais complexos, mais ativos e mais diversos.

Hoje em dia, a fotossíntese e o metabolismo aeróbico são partes complementares de um grande ciclo de energia que engloba todo o planeta (Fig. 19.1). Por meio da fotossíntese, dióxido de carbono e água são usados para a síntese de materiais orgânicos e, por meio do metabolismo oxidativo, materiais orgânicos são decompostos em dióxido de carbono e água. Em eucariontes, estes processos ocorrem em organelas subcelulares especializadas – a fotossíntese nos **cloroplastos** e o metabolismo aeróbico nas **mitocôndrias**. Ambas estas organelas parecem ter surgido de organismos procarióticos simples que se estabeleceram dentro das células eucarióticas, um fenômeno denominado **endossimbiose**, há mais de um bilhão de anos. Como tais organismos trouxeram seus genomas, bem como suas fontes fotossintéticas e oxidativas, hoje todos os cloroplastos e mitocôndrias eucarióticos contêm DNA. Neste capítulo, exploraremos a genética e biologia molecular destas importantes organelas.

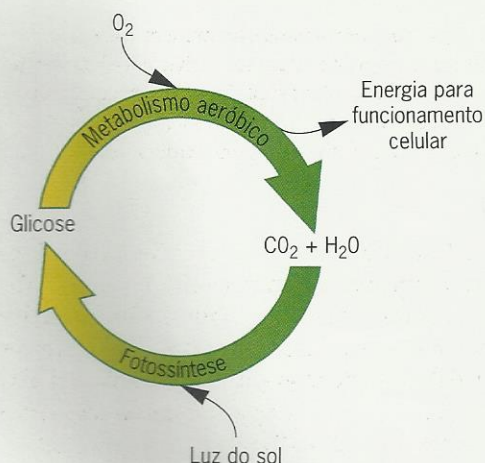


Fig. 19.1 ■ A engrenagem energética biológica.

## A GENÉTICA CLÁSSICA DAS ORGANELAS

O material genético das organelas foi estudado seguindo-se a herança de pigmentação de folhas em plantas, a resistência a antibióticos em algas e a competência metabólica em leveduras.

A descoberta do DNA de cloroplastos e mitocondrial foi antecipada por estudos que indicaram a existência de fatores hereditários fora do núcleo. Tais estudos datam do começo do século vinte. De fato, o primeiro deles foi feito por Carl Correns, um dos três botânicos que redescobriram os princípios de Mendel.

A maioria dos estudos iniciais foram feitos com plantas, que possuem tanto mitocôndrias quanto cloroplastos. Conseqüentemente, nem sempre foi possível determinar qual

dos dois tipos de organelas era responsável pela herança não nuclear. Estudos posteriores usaram leveduras, para as quais o envolvimento de cloroplastos podia ser excluído. Nas seções seguintes, faremos uma revisão de alguns destes estudos clássicos. O princípio comum compartilhado por todos eles é que a herança das organelas não segue regras mendelianas simples. Em vez disso, ela é caracterizada por contribuições desiguais dos dois genitores e por uma segregação irregular dos alelos, isto é, viola o princípio mendeliano da segregação. A herança das organelas é portanto uma *hereditariedade não mendeliana*.

## Variação de Folhas em Plantas

As folhas em plantas às vezes exibem um padrão marcante de **variação de cor** (Fig. 19.2). Alguns setores são verdes,

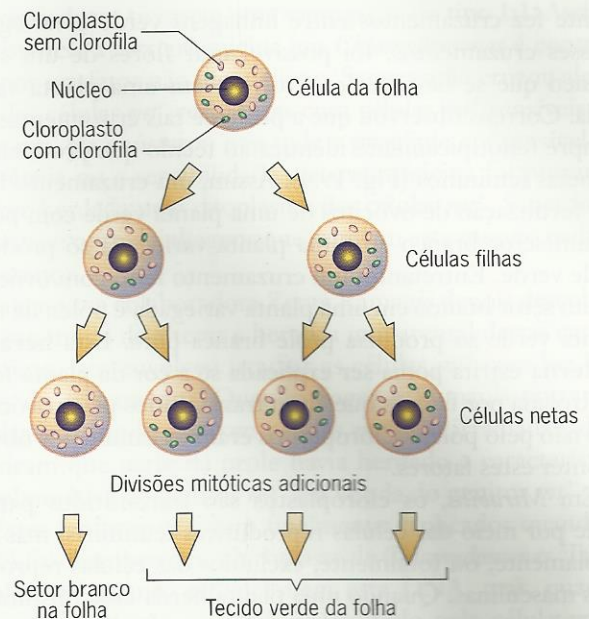




**Fig. 19.2** ■ Variação de folhas causada pela segregação de tipos diferentes de cloroplastos.

outros são brancos. Também podem existir tons intermediários de verde ou amarelo-claro. Esses padrões em mosaico são altamente apreciados por suas qualidades ornamentais, o que explica parcialmente por que tantos pesquisadores se interessaram em estudá-los.

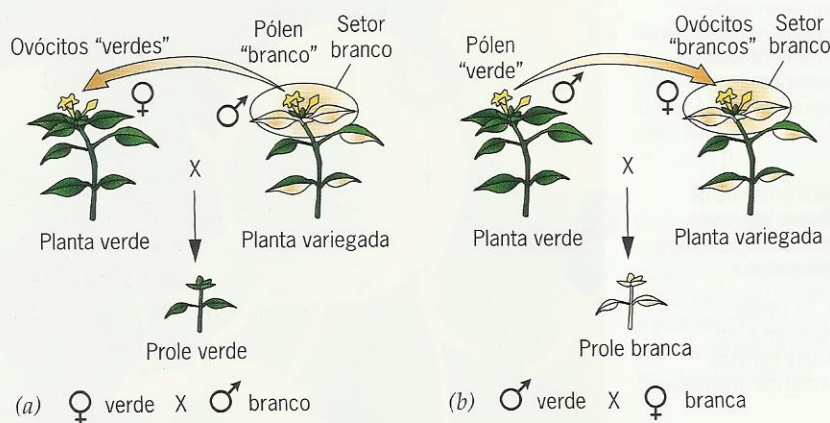
Muitos casos de variação de folhas podem ser explicados pela distribuição de tipos diferentes de cloroplastos durante divisão celular (Fig. 19.3). As células podem conter alguns cloroplastos que são capazes de produzir o pigmento verde clorofila e outros que não são. Esta mistura de dois tipos de organelas dentro de uma célula é chamada de **heteroplasmia**, das palavras gregas que significam “dois corpos diferentes”. A presença de um único tipo de organela dentro das células é chamada de **homoplasmia**. Quando uma célula heteroplásmica se divide, os dois tipos de cloroplastos são distribuídos para as células filhas de modo irregular. Ao longo de várias divisões, essa distribuição irregular pode produzir uma célula que não tem cloroplastos produtores de pigmentos. Tal célula então prolifera em um setor branco em uma



**Fig. 19.3** ■ Distribuição de cloroplastos durante a mitose. A distribuição irregular de cloroplastos durante a divisão celular pode produzir uma célula que não possui cloroplastos capazes de fazer clorofila. Por meio de outras divisões, tal célula irá produzir um setor branco de tecido em uma folha de outro modo verde.

folha de outro modo verde. Os setores brancos produzidos por este processo são irregulares em tamanho, forma e posição. (Em algumas plantas, os setores brancos são formados em padrões previsíveis e definidos nas folhas. Estes setores não se devem a distribuição de cloroplastos. Eles são criados por sistemas fisiológicos que regulam a síntese do pigmento em um arcabouço espacial e temporal dentro das folhas.)

O efeito da variação de folhas foi primeiro estudado pelos botânicos alemães Carl Correns e Erwin Baur. Correns trabalhou com linhagens variegadas da maravilha, *Mirabilis jalapa*, uma planta de jardim muito popular. Ele sistematicamente



**1** Correns fez cruzamentos recíprocos entre plantas *Mirabilis* verdes e variegadas, usando setores brancos nas plantas variegadas como fonte de gametas.

**2** De cada cruzamento, cada prole apresentou o mesmo fenótipo que o genitor feminino.

Logo, na planta *Mirabilis* a cor da planta exibe estrita herança materna.

**Fig. 19.4** ■ Experimentos de Correns sobre a herança de variação de folhas em *Mirabilis*. (a) Fêmea verde × macho branco (em uma planta variegada). (b) Macho verde × fêmea branca (em uma planta variegada).

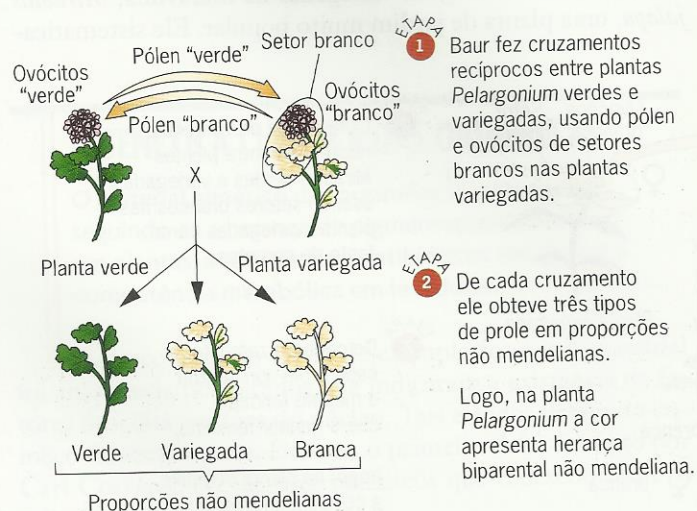


mente fez cruzamentos entre linhagens verde e variegada. Nesses cruzamentos, foi possível usar flores de um setor branco que se haviam desenvolvido em uma planta variegada. Correns observou que a prole de tais cruzamentos era sempre fenotipicamente idêntica ao tecido que produziu os gametas femininos (Fig. 19.4). Assim, um cruzamento feito por fertilização de ovócitos de uma planta verde com pólen de um setor branco em uma planta variegada só produzia prole verde. Entretanto, um cruzamento feito com ovócitos de um setor branco em uma planta variegada e pólen de uma planta verde só produzia prole branca pura. Esta **herança materna** estrita podia ser explicada se a cor da planta fosse controlada por fatores que eram transmitidos pelos ovócitos mas não pelo pólen. Cloroplastos eram os candidatos óbvios a conter estes fatores.

Em *Mirabilis*, os cloroplastos são transmitidos para a prole por meio das células reprodutivas femininas, mas são amplamente, ou totalmente, excluídos das células reprodutivas masculinas. Quando uma planta herda uma mistura de cloroplastos pigmentados e não pigmentados do genitor feminino, seus tecidos podem variegar porque os dois tipos de cloroplastos distribuem-se durante o desenvolvimento.

Baur estudou a variação em outra espécie ornamental, *Pelargonium zonale* (Fig. 19.5). Nesta planta, cruzamentos entre as linhagens verde e variegada produzem uma mistura de prole, algumas verdes, algumas variegadas e algumas brancas puras. Entretanto, tais fenótipos diferentes não são recuperados em proporções mendelianas. Esta **herança biparental não mendeliana** indica que a cor das folhas de *Pelargonium* é determinada por uma mistura de fatores maternos e paternos situados fora do núcleo, supostamente nos cloroplastos. Portanto, parece que em *Pelargonium* os cloroplastos são transmitidos pelo pólen bem como pelo ovócito.

Os experimentos de Correns e Baur demonstraram que a variação de folhas é herdada como uma característica

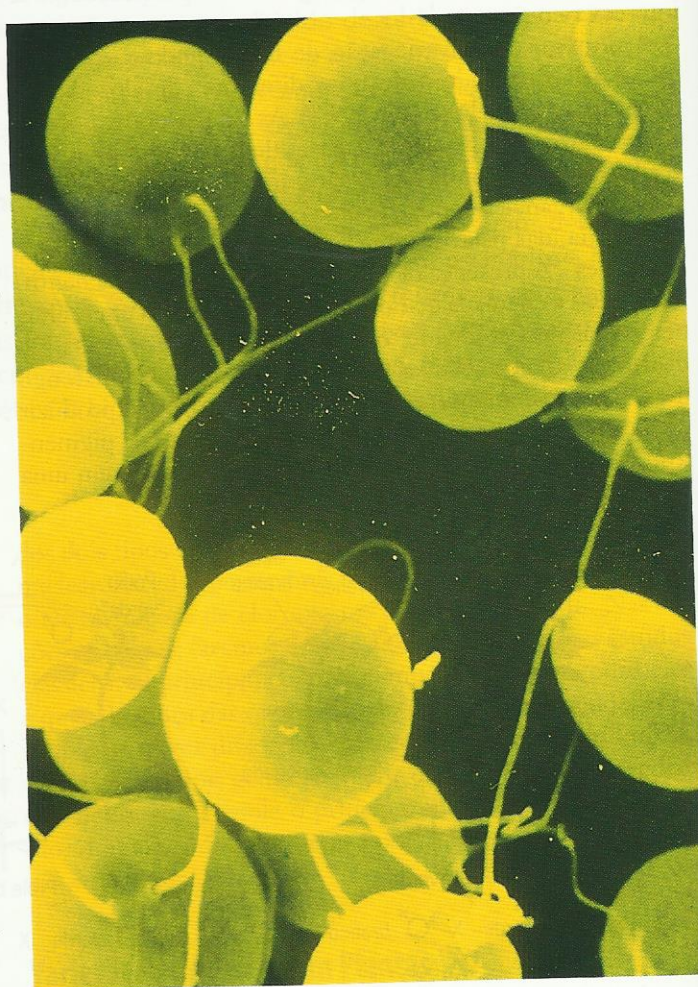


**Fig. 19.5** ■ Experimentos de Baur sobre a herança de variação de folhas em *Pelargonium*. Ambos os cruzamentos recíprocos produzem três tipos de prole em proporções não mendelianas.

não mendeliana, aquela que é quase certamente controlada por fatores situados nos cloroplastos. A descoberta de que os cloroplastos contêm DNA tornou esta hipótese muito atrativa. Mas, a despeito do trabalho de muitos cientistas, a base molecular da variação de folhas ainda é desconhecida. Provavelmente ela envolve diferenças nas moléculas de DNA que estão presentes dentro de cloroplastos normais e os deficientes de pigmento. É necessário mais trabalho para determinar a causa deste interessante fenômeno botânico.

### Resistência a Antibióticos em *Chlamydomonas*

Algas verdes fotossintéticas são uma parte importante da biosfera. Elas são abundantes nos oceanos, onde formam a base da cadeia alimentar, mas também são encontradas em terra. Uma espécie terrestre, *Chlamydomonas reinhardtii* (Fig. 19.6), foi amplamente usada em pesquisas genéticas. Este organismo unicelular haplóide existe em dois tipos reprodutivos diferentes, representados por mais e menos. Células de tipos reprodutivos diferentes podem fundir-se para produzir um zigoto diplóide. O zigoto então sofre meiose para



**Fig. 19.6** ■ Células da alga unicelular *Chlamydomonas reinhardtii*. Fonte: R. Kessel & C. R. Shih, Visuals Unlimited.



produzir quatro proles haplóides, que se dividem mitoticamente para produzir clones de células vegetativas. Análise destes clones indica que o tipo reprodutivo é controlado por um gene nuclear com dois alelos,  $mt^+$  (especificando o tipo reprodutivo mais) e  $mt^-$  (especificando o tipo reprodutivo menos). Duas das quatro proles de um zigoto de *Chlamydomonas* são portanto  $mt^+$  e duas são  $mt^-$ . Muitos outros genes nucleares com um padrão mendeliano de segregação foram identificados em *Chlamydomonas*.

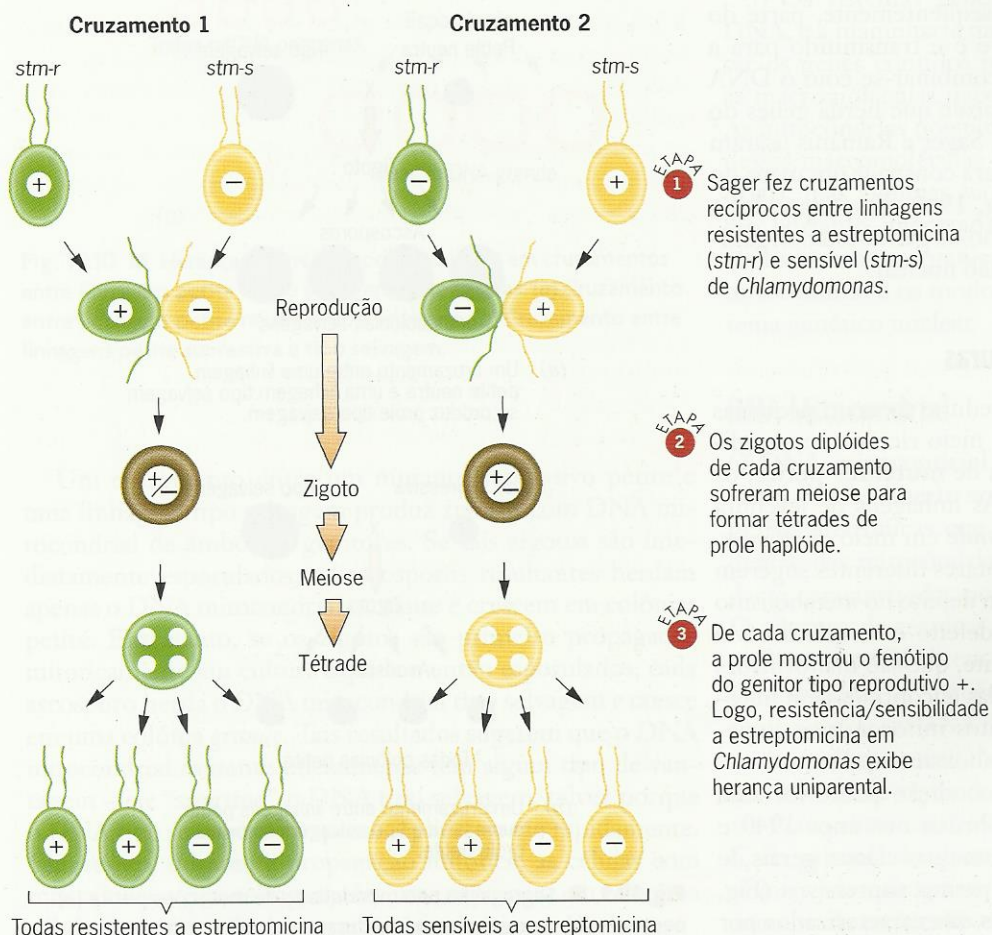
Cada célula haplóide de *Chlamydomonas* contém um único cloroplasto grande e várias mitocôndrias. Quando células  $mt^+$  e  $mt^-$  se fundem, suas organelas são combinadas na mesma célula diplóide. Não está claro o que ocorre com estas organelas dentro do zigoto, mas, quando são produzidas as quatro haplóides da prole, cada uma tem um cloroplasto e várias mitocôndrias. O número básico de organelas é portanto preservado.

Em 1954, Ruth Sager, uma geneticista americana, descobriu que a resistência a antibióticos em *Chlamydomonas* é herdada como uma característica não mendeliana (Fig. 19.7). Sager cruzou uma linhagem mutante  $mt^+$  que era resistente ao antibiótico estreptomicina com uma linhagem  $mt^-$  que era sensível a este antibiótico. Toda a prole desse cruzamento era resistente a estreptomicina, muito embora os alelos do

tipo reprodutor tivessem uma segregação do tipo 1:1. Assim, a resistência a estreptomicina em *Chlamydomonas* é controlada por um fator não mendeliano. Sager então cruzou algumas das células  $mt^-$  resistentes com células  $mt^+$  sensíveis, e, desta vez, ela constatou que toda a progênie era sensível. A resistência ou a sensibilidade a estreptomicina foi portanto sempre herdada pelo citoplasma das células  $mt^+$ . Sager logo encontrou outras linhagens que seguiam este mesmo padrão de herança.

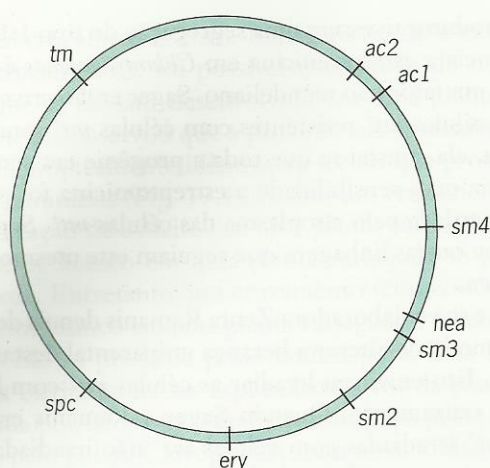
Sager e sua colaboradora Zenta Ramanis depois descobriram um modo de alterar a herança uniparental destas características. Isto envolveu irradiar as células  $mt^+$  com luz UV antes do cruzamento. Quando Sager e Ramanis cruzaram células  $mt^+$  irradiadas com células  $mt^-$  não irradiadas, descobriram que parte da prole havia herdado a característica citoplasmática do genitor  $mt^-$  em vez da do genitor  $mt^+$ .

Estes fenômenos foram finalmente explicados quando o DNA foi descoberto no cloroplasto de *Chlamydomonas*. Tanto células  $mt^+$  quanto  $mt^-$  possuem este DNA, mas, quando ocorre reprodução, o DNA contribuído pela célula  $mt^-$  é degradado, provavelmente por uma nuclease que especificamente o reconhece. A prole de um cruzamento portanto só conserva o DNA herdado do cloroplasto  $mt^+$ . Como resultado, as características codificadas por este DNA – e a resis-



**Fig. 19.7** ■ Experimentos de Sager mostrando herança uniparental de resistência ( $stm-r$ ) e sensibilidade ( $stm-s$ ) a estreptomicina em *Chlamydomonas*. O tipo reprodutivo é indicado por um mais (+) ou menos (-). Toda a prole de um cruzamento tem o mesmo fenótipo do genitor  $mt^+$ .





**Fig. 19.8** ■ Um mapa genético do DNA de cloroplasto em *Chlamydomonas*, baseado no trabalho de Sager e Ramanis. Símbolos: *ac1*, *ac2*, necessidade de acetato; *sm4*, dependência de estreptomicina; *nea*, resistência a neamina; *sm3*, resistência a baixo nível de estreptomicina; *sm2*, resistência a alto nível de estreptomicina; *ery*, resistência a eritromicina; *spc*, resistência a espectinomicina; *tm*, sensibilidade a temperatura.

tência a estreptomicina é uma delas – são sempre herdadas do genitor *mt*<sup>+</sup>. Quando células *mt*<sup>+</sup> são irradiadas antes da reprodução, o sistema que destrói o DNA *mt*<sup>-</sup> do cloroplasto é parcialmente incapacitado. Conseqüentemente, parte do DNA do cloroplasto *mt*<sup>-</sup> sobrevive e é transmitido para a prole. Esse DNA pode de fato recombinar-se com o DNA do cloroplasto *mt*<sup>+</sup>, gerando uma prole que herda genes do cloroplasto de ambos os genitores. Sager e Ramanis usaram a frequência de tal recombinação para construir um mapa do “cromossomo” do cloroplasto (Fig. 19.8). Esse mapa foi a primeira imagem detalhada de como os genes estão organizados em uma molécula de DNA não nuclear.

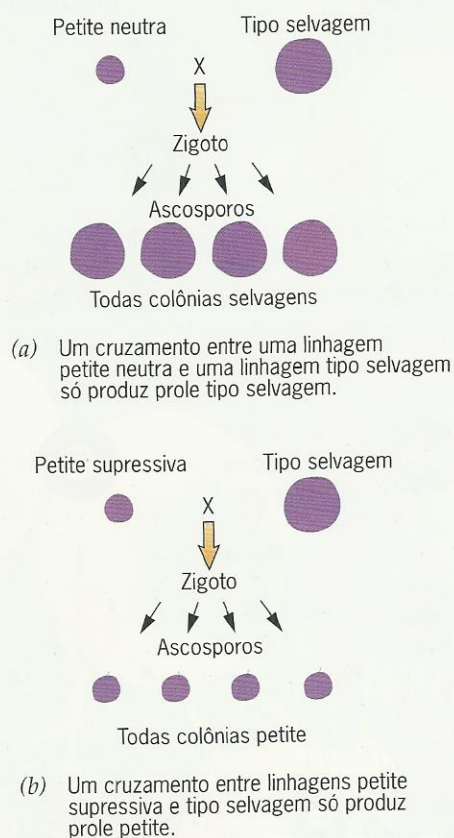
### Defeitos Metabólicos em Leveduras

Algumas linhagens mutantes de leveduras formam pequenas colônias quando cultivadas em um meio rico contendo glicose. Essas linhagens são chamadas de **mutantes petite**, da palavra francesa para “pequeno”. As linhagens de levedura tipo selvagem formam colônias **grande** em meio de glicose. Testes adicionais com fontes alimentares diferentes sugerem que os mutantes petite sofrem de um defeito no metabolismo da glicose. Em muitos casos, este defeito é correlacionável com as mitocôndrias das células petite, que são malformadas e não têm muitas das macromoléculas encontradas em mitocôndrias tipo selvagem. Estes defeitos mitocondriais impedem que células petite efetuem metabolismo aeróbico.

A primeira análise genética da condição petite foi feita pelo pesquisador francês Boris Ephrussi nos anos 1940 e 1950. Sua análise revelou que existem duas classes gerais de mutantes petite, petites neutros e petites supressivos (Fig. 19.9). Os **mutantes petite neutros** são caracterizados por

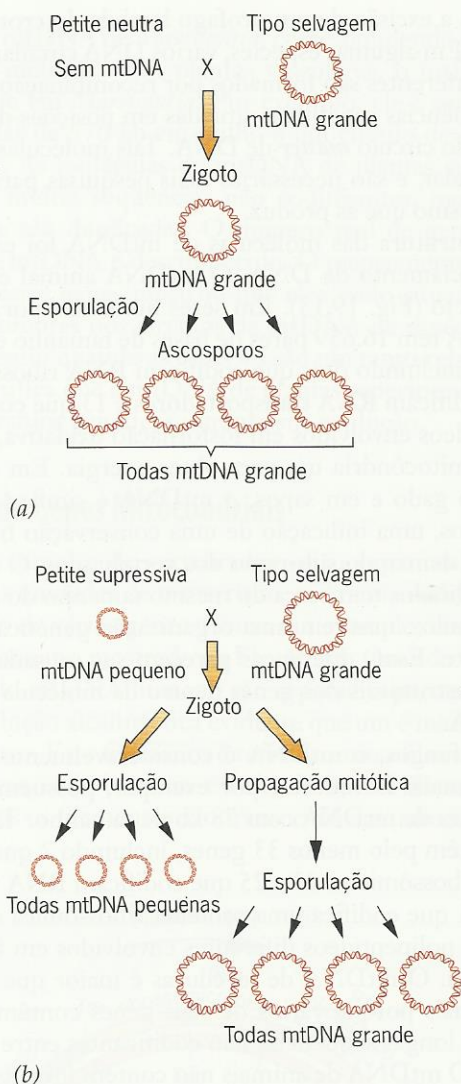
uma incapacidade de transmitir o fenótipo petite para a prole de cruzamentos com linhagens do tipo selvagem. Quando são feitos tais cruzamentos, todas as quatro proles haplóides de um par de células cruzadas crescem em grandes colônias, sugerindo que a mutação petite foi perdida. Em contraste, **mutantes petite supressivos** são, em condições apropriadas, capazes de transmitir o fenótipo petite para toda sua prole, sugerindo que a condição tipo selvagem foi perdida. Com ambos os tipos de mutantes, cruzamentos recíprocos entre linhagens tipo mutante e tipo selvagem dão resultados idênticos. Assim, os tipos reprodutivos de leveduras não estão envolvidos na herança da característica petite.

A base molecular dos dois tipos de mutantes petite foi determinada nos anos 1970 e 1980. Petites neutros não têm nenhum DNA mitocondrial, e petites supressivos têm um DNA mitocondrial muito mutado (Fig. 19.10). Um cruzamento entre um mutante petite neutro e um tipo selvagem produz zigotos que herdam mitocôndrias de ambos os genitores. Entretanto, apenas as mitocôndrias do genitor tipo selvagem contêm DNA. Quando tais zigotos passam pela meiose e formam esporos, o DNA mitocondrial tipo selvagem é distribuído para cada um dos quatro ascósporos. À medida que eles crescem e se dividem, desenvolvem-se mitocôndrias funcionais e saudáveis, permitindo que as células façam metabolismo aeróbico. Assim, quando são formadas colônias, todas elas têm o fenótipo *grande*.



**Fig. 19.9** ■ Segregação não mendeliana de mutações petite (a) neutras e (b) supressivas em leveduras.





**Fig. 19.10** ■ Herança de DNA mitocondrial (mt) em cruzamentos entre linhagens *petite* e tipo selvagem de leveduras. (a) Cruzamento entre linhagens *petite neutra* e tipo selvagem. (b) Cruzamento entre linhagens *petite supressiva* e tipo selvagem.

Um cruzamento entre um mutante supressivo *petite* e uma linhagem tipo selvagem produz zigotos com DNA mitocondrial de ambos os genitores. Se tais zigotos são imediatamente esporulados, os ascósporos resultantes herdam apenas o DNA mitocondrial mutante e crescem em colônias *petite*. Entretanto, se os zigotos são primeiro propagados mitoticamente em cultura líquida e então esporulados, cada ascósporo herda o DNA mitocondrial tipo selvagem e cresce em uma colônia *grande*. Tais resultados sugerem que o DNA mitocondrial mutante inicialmente tem algum tipo de vantagem – ele “suprime” o DNA tipo selvagem, talvez porque o DNA tipo selvagem não possa replicar-se tão rapidamente. Entretanto, durante a propagação mitótica, as células com mitocôndrias tipo selvagem têm uma vantagem em relação a células com mitocôndrias mutantes e dominam a cultura.

Qual a natureza do DNA mitocondrial em mutantes *petite* supressivos? Análises físicas e químicas mostraram que ele é uma molécula de DNA menor que o DNA mitocondrial tipo selvagem e também é rico em A:T. Estas duas características podem dotar o DNA mitocondrial *petite* supressivo de uma vantagem replicativa em relação ao tipo selvagem. Entretanto, uma compreensão total do motivo pelo qual o DNA *petite* supressivo é tão efetivamente transmitido através de cruzamentos exigirá mais pesquisas.

### PONTOS IMPORTANTES

- A hereditariedade de organelas é caracterizada por contriuições parentais desiguais e por segregações fenotípicas irregulares.
- As características não mendelianas da hereditariedade de organelas se devem, em parte, à transmissão preferencial de cloroplastos e mitocôndrias através dos gametas de um sexo, que em eucariontes superiores geralmente é o feminino.

## A GENÉTICA MOLECULAR DE MITOCÔNDRIAS

As mitocôndrias contêm moléculas de DNA com uma variedade de genes.

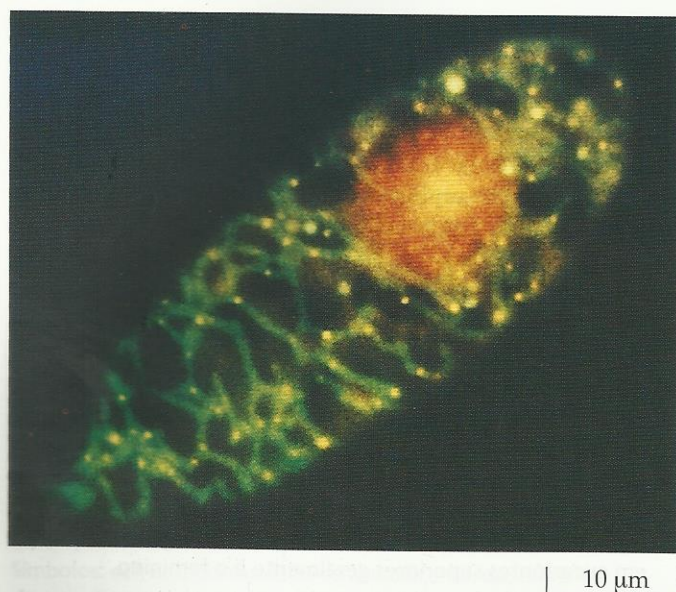
Os sistemas genéticos mitocondriais consistem em DNA, e a maquinaria molecular precisa replicar-se e expressar os genes contidos nesse DNA. Essa maquinaria inclui as macromoléculas necessárias para transcrição e tradução. As mitocôndrias possuem seus próprios ribossomos. Muitas dessas macromoléculas são codificadas por genes mitocondriais, mas algumas são codificadas por genes nucleares e são portanto importadas do citosol. Nas seções seguintes, exploraremos a organização e a função do sistema genético mitocondrial e os modos pelos quais ele interage com o sistema genético nuclear.

### DNA Mitocondrial

O **DNA mitocondrial**, ou **mtDNA** como às vezes é abreviado, foi descoberto nos anos 1960, inicialmente por micrografias eletrônicas que revelaram fibras similares ao DNA dentro das mitocôndrias. Depois, essas fibras foram extraídas e caracterizadas por procedimentos físicos e químicos. O advento de técnicas de DNA recombinante possibilitou analisar o mtDNA em maiores detalhes. De fato, já foram determinadas as seqüências completas de nucleotídeos das moléculas de mtDNA de muitas espécies diferentes.

As moléculas de DNA mitocondrial variam muito de tamanho, desde cerca de 6 kb no parasita causador da malária *Plasmodium* até 2.500 kb em algumas plantas com flores. Cada mitocôndria parece conter várias cópias de DNA, e, como cada célula tem geralmente muitas mitocôndrias, o número de moléculas de mtDNA por célula pode ser muito





**Fig. 19.11** ■ DNA mitocondrial (amarelo) no organismo unicelular *Euglena gracilis*. O DNA nuclear (vermelho) é também visível.

grande (Fig. 19.11). Em um ovócito de vertebrado, por exemplo, foi estimado que estão presentes até  $10^8$  cópias de mtDNA. As células somáticas, entretanto, têm menos cópias, talvez menos de 1.000.

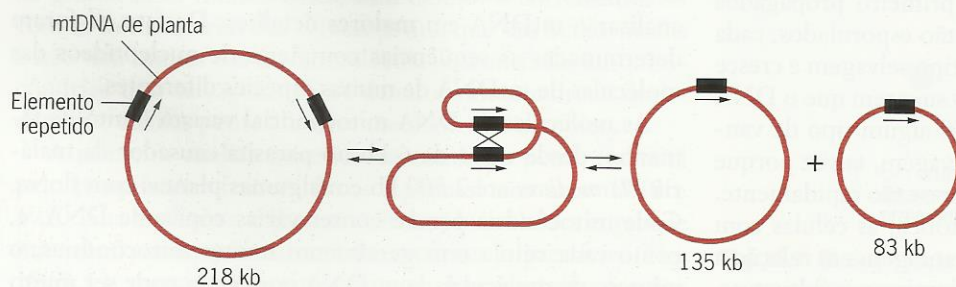
A maioria das moléculas de mtDNA é circular, mas em algumas espécies, tais como a alga *Chlamydomonas reinhardtii* e o ciliado *Paramecium aurelia*, elas são lineares. As moléculas circulares de mtDNA, que foram as mais completamente estudadas, parecem estar organizadas de muitos modos diferentes. Nos vertebrados, 37 genes distintos estão embalados em um círculo com 16 a 17 kb, deixando pouco ou nenhum espaço entre os genes. Em algumas das plantas com flores, um número desconhecido de genes está disperso em uma grande molécula circular de DNA com centenas ou milhares de quilobases de tamanho. Nestas plantas, os genes mitocondriais podem ser separados em moléculas circulares diferentes por um processo de recombinação intramolecular (Fig. 19.12). Essa recombinação é mediada por seqüências repetitivas situadas no mtDNA. Uma troca entre duas das seqüências repetitivas pode separar o círculo de mtDNA "master" em dois círculos menores, um processo que superficialmente

lembra a excisão de um profago lambda do cromossomo de *E. coli*. Em algumas espécies, vários DNA circulares de tamanhos diferentes são formados por recombinação entre pares de seqüências repetitivas situadas em posições diferentes ao redor do círculo master de DNA. Tais moléculas são difíceis de estudar, e são necessárias mais pesquisas para elucidar o mecanismo que as produz.

A estrutura das moléculas de mtDNA foi estudada por seqüenciamento de DNA. O mtDNA animal é pequeno e compacto (Fig. 19.13). Em seres humanos, por exemplo, o mtDNA tem 16.659 pares de bases de tamanho e contém 37 genes, incluindo dois que codificam RNA ribossômicos, 22 que codificam RNA transportadores e 13 que codificam polipeptídeos envolvidos em fosforilação oxidativa, o processo que a mitocôndria usa para obter energia. Em camundongos, no gado e em sapos, o mtDNA é similar ao de seres humanos, uma indicação de uma conservação básica de estrutura dentro do sub-ramo dos vertebrados. O mtDNA de invertebrados tem cerca do mesmo tamanho do mtDNA de vertebrados, mas tem uma organização genética um pouco diferente. Essas diferenças parecem ser causadas por rearranjos estruturais dos genes dentro da molécula circular de mtDNA.

Em fungos, o mtDNA é consideravelmente maior que em animais. Leveduras, por exemplo, possuem moléculas circulares de mtDNA com 78 kb de tamanho. Tais moléculas contêm pelo menos 33 genes, incluindo 2 que codificam RNA ribossômicos, 23 a 25 que codificam RNA transportadores, 1 que codifica uma proteína ribossômica e 7 que codificam polipeptídeos diferentes envolvidos em fosforilação oxidativa. O mtDNA de leveduras é maior que o mtDNA de animais porque vários de seus genes contêm íntrons, e existem longas seqüências não codificantes entre alguns dos genes. O mtDNA de animais não contém íntrons.

O mtDNA de plantas é muito maior que o mtDNA de outros organismos. Ele é também de estrutura mais variável. Estas conclusões vieram de análises físicas e químicas e de seqüenciamento do DNA. Um dos primeiros mtDNA de plantas a ser seqüenciado é de uma hepática, *Marchantia polymorpha*. O mtDNA desta planta primitiva não vascular é uma molécula circular com 186 kb e 94 matrizes de leitura aberta (ORF), algumas correspondendo a genes conhecidos e outras sem funções genéticas conhecidas. Estas últimas ORF são portanto chamadas de URF, de matrizes de leitura sem atribuições (*unassigned*). Foram encontrados 32 íntrons



**Fig. 19.12** ■ Recombinação intramolecular no mtDNA de couve chinesa, *Brassica campestris*. A recombinação entre os elementos repetidos na grande molécula circular de DNA divide esta molécula em duas menores. Alternativamente, os elementos repetidos nas duas moléculas menores podem recombinar-se para produzir uma única molécula grande.



distintos no mtDNA de *Marchantia*, correspondendo a cerca de 20% da molécula. Em plantas vasculares, o mtDNA é maior do que em *Marchantia*. Por exemplo, ele é uma molécula circular de 570 kb em milho e um círculo de 300 kb na melancia. As moléculas de mtDNA de plantas superiores contêm muitas seqüências não codificantes, incluindo algumas que são duplicadas. O número real de genes por molécula de mtDNA é desconhecido. O mapeamento físico de alguns destes genes mostrou que eles estão situados em posições diferentes nos círculos de mtDNA de espécies diferentes, mesmo quando as espécies são um tanto relacionadas. Isto significa que o mtDNA de plantas superiores sofreu muitos rearranjos genéticos durante sua evolução.

### Expressão de Genes Mitocondriais

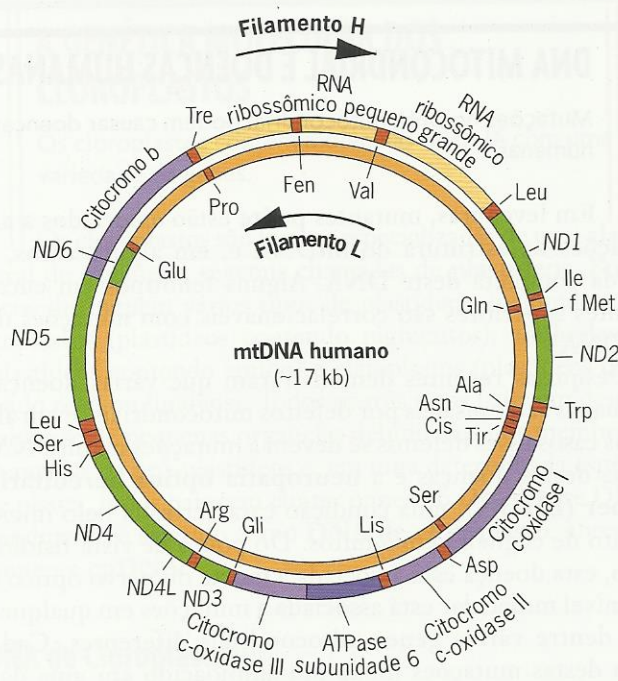
Os mtDNA simples de vertebrados estão organizados em duas grandes unidades de transcrição, cada uma codificando a informação de vários genes. Como um exemplo, considere o mtDNA humano mostrado na Fig. 19.13. Quando os dois filamentos do mtDNA humano são separados por centrifugação em solução alcalina, fica evidente que um é mais denso do que o outro. Este filamento é referido como filamento H (pesado) e seu complemento é referido como filamento L (leve). Ambos os filamentos são transcritos em mitocôndrias

humanas. Os promotores para as unidades de transcrição H e L estão situados antecedendo o gene de tRNA de fenilalanina. Os transcritos que iniciam nestes pontos são ampliados em direções opostas ao redor de toda a circunferência da molécula de mtDNA. O transcrito do filamento H codifica os dois RNA ribossômicos, 14 tRNA e 12 polipeptídeos e o transcrito do filamento L codifica 8 tRNA e 1 polipeptídeo. Cada transcrito é clivado para separar os tRNA dos rRNA e mRNA, e os mRNA são poliadenilados. Cada mRNA é então traduzido em polipeptídeos, usando ribossomos mitocondriais e uma combinação de tRNA nucleares e ribossômicos.

A tradução nas mitocôndrias ocorre como nos ribossomos do citosol, exceto que alguns dos códons têm um significado diferente. AGA e AGG são códons de fim nas mitocôndrias de mamíferos, enquanto no citosol eles especificam a incorporação de arginina em um polipeptídeo. UGA, que é um códon de fim no citosol, é um códon de triptofano nas mitocôndrias; e AUA, que codifica isoleucina no citosol, é o códon metionina de iniciação nas mitocôndrias. Estas e outras variações de códons mitocondriais indicam que o código genético não é totalmente universal. Evidentemente, as mitocôndrias desenvolveram sua própria variação do código genético, supostamente como consequência de sua evolução a partir de organismos de vida livre que foram incorporados a células eucarióticas há mais de um bilhão de anos. Voltaremos ao assunto da evolução mitocondrial no final deste capítulo.

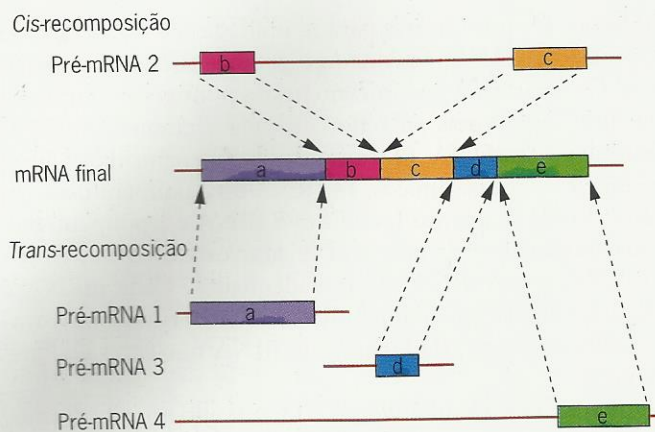
Em fungos e plantas, o mtDNA é organizado em muitas unidades de transcrição separadas, algumas contendo a informação para mais de um gene. Sabemos pouco sobre os detalhes da transcrição, mas em leveduras a RNA-polimerase mitocondrial é um único polipeptídeo codificado por um gene nuclear. O processamento do RNA separa transcritos mitocondriais de planta em suas partes constituintes e também remove os íntrons, que estão presentes em vários genes mitocondriais de plantas. No momento, a mecânica destes eventos é pouco compreendida.

Outra peculiaridade da expressão gênica mitocondrial é que muitos dos transcritos de mtRNA sofrem **edição**, isto é, alguns dos nucleotídeos são mudados após a síntese do transcrito. A mudança mais freqüente é de C para U, mas ocasionalmente U é mudado para C. Assim, a edição do RNA altera a composição de códons em transcritos mitocondriais da planta, incluindo alguns que indicariam o fim de uma síntese de polipeptídeos. A edição altera a informação que é codificada no mtDNA e permite que sejam sintetizados polipeptídeos funcionais. Curiosamente, a edição do RNA mitocondrial não é encontrada nas plantas não vasculares (musgos e algas), muito embora todos os grupos de plantas superiores (samambaias, gimnospermas e angiospermas) pareçam tê-la. Assim, o mecanismo de edição provavelmente evoluiu subsequentemente ao estabelecimento das plantas na terra. Edição do RNA também ocorre nas mitocôndrias de protozoários, incluindo tripanossomas, em que os mecanismos foram estudados em mais detalhes. Nestes orga-



**Fig. 19.13** ■ Mapa do mtDNA humano mostrando o padrão de transcrição. Os genes no círculo interno são transcritos do filamento L do DNA, enquanto os genes do círculo externo são transcritos do filamento H do DNA. As setas mostram o sentido da transcrição. ND1–6 são genes que codificam subunidades da enzima NADH-redutase; os genes de tRNA no mtDNA são indicados por abreviações para os aminoácidos.





**Fig. 19.14** ■ Trans-recomposição em mitocôndria de trigo. Quatro RNA diferentes contribuem para o mRNA final que codifica um polipeptídeo da enzima NADH-redutase.

nismos, pequenas moléculas de RNA que são parcialmente complementares aos mtDNA transcritos servem como guias para o processo de edição (veja Fig. 11.22). Eles são portanto chamados de *RNA guias* (gRNA). Um mecanismo similar de guia pode operar em plantas, mas os detalhes são atualmente desconhecidos.

Uma terceira peculiaridade da expressão gênica mitocondrial é que alguns RNA mensageiros mitocondriais são formados pelo processo de *trans-splicing* (trans-recomposição). O *trans-splicing* ocorre quando segmentos de um gene são dispersos na molécula de mtDNA. Cada segmento do gene é transcrito independentemente, e os éxons de transcritos diferentes são recompostos por interações dos íntrons que os flanqueiam. Por exemplo, em trigo, o gene *nad1*, que codifica uma subunidade de NADH-redutase, uma proteína de fosforilação oxidativa, é partido em quatro segmentos no mtDNA. Cada um destes segmentos é transcrito separadamente, e os transcritos resultantes são então recompostos para formar o mRNA (Fig. 19.14). Este processo requer uma única reação de *cis-recomposição* e três reações de *trans-recomposição*.

### Intercâmbio Entre Produtos de Genes Mitocondriais e Nucleares

A maioria – e talvez a totalidade – dos produtos gênicos mitocondriais só funcionam dentro das mitocôndrias. Entretanto, eles não funcionam sozinhos. Muitos produtos gênicos nucleares são importados para aumentar ou facilitar seu funcionamento. Os ribossomos mitocondriais, por exemplo, são construídos com RNA ribossômico transcrito de genes mitocondriais e com proteínas ribossômicas codificadas por genes nucleares. As proteínas ribossômicas são sintetizadas no citosol e importadas para as mitocôndrias para montagem em ribossomos.

Muitos dos polipeptídeos necessários para o metabolismo aeróbico também são produzidos no citosol. Eles

incluem subunidades de várias proteínas envolvidas em fosforilação oxidativa – por exemplo, a ATPase que é responsável pela ligação de energia do metabolismo aeróbico em ATP. Entretanto, como algumas das subunidades desta proteína são sintetizadas nas mitocôndrias, a proteína completa é de fato uma mistura de produtos gênicos nucleares e mitocondriais. Esta composição dupla sugere que os sistemas genéticos nuclear e mitocondrial sejam coordenados de algum modo para que sejam feitas quantidades equivalentes de seus produtos. Atualmente estão sendo investigados possíveis mecanismos moleculares para esta coordenação.

### PONTOS IMPORTANTES

- As moléculas de DNA mitocondrial (mtDNA) variam de 6 kb a 2.500 kb de tamanho, e a maioria parece ser circular.
- As moléculas de DNA mitocondrial contêm genes para alguns dos RNA ribossômicos, RNA transportadores e polipeptídeos usados dentro da mitocôndria.
- A estrutura, a organização e a expressão de genes mitocondriais variam entre espécies.
- Em alguns organismos, os transcritos de genes mitocondriais são editados após serem sintetizados.
- São necessários tanto produtos gênicos mitocondriais quanto nucleares para funcionamento mitocondrial normal.

## DNA MITOCONDRIAL E DOENÇAS HUMANAS

Mutações no DNA mitocondrial podem causar doenças humanas.

Em leveduras, mutantes *petite* estão associados a alterações na estrutura do mtDNA e, em alguns casos, à perda completa deste DNA. Alguns fenótipos em eucariotes superiores são correlacionáveis com mutações de mtDNA?

Pesquisas recentes demonstraram que várias doenças humanas são causadas por defeitos mitocondriais, e em alguns casos esses defeitos se devem a mutações no mtDNA. Uma destas doenças é a **neuropatia óptica hereditária Leber (LHON)**, uma condição caracterizada pelo início súbito de cegueira em adultos. Do ponto de vista fisiológico, esta doença está associada a morte do nervo óptico e em nível molecular está associada a mutações em qualquer um dentre vários genes mitocondriais diferentes. Cada uma destas mutações muda um aminoácido em uma das proteínas mitocondriais, reduzindo assim a eficiência da fosforilação oxidativa. A redução é grande o suficiente para destruir o funcionamento do nervo óptico e causar cegueira total. Não sabemos por que este efeito letal é limitado ao nervo óptico. Talvez as células nervosas sejam especialmente sensíveis a perturbações no metabolismo aeróbico. Como seria esperado para uma mutação



mitocondrial, LHON é herdada estritamente através da linhagem materna. Nunca há transmissão pelos homens. Assim, a prole de uma mulher doente pode expressar o fenótipo da doença, enquanto a prole de um homem doente não.

Outro distúrbio causado por uma mutação no mtDNA é a **síndrome medula-pâncreas Pearson**. Esta doença, caracterizada por perda das células da medula óssea durante a infância, é freqüentemente fatal. Ela é causada por uma deleção relativamente grande no mtDNA. Pessoas com a síndrome Pearson quase nunca têm genitores afetados. Assim, a deleção causadora provavelmente ocorre espontaneamente durante o desenvolvimento da criança ou na ovocitogênese da mãe. Pessoas com a síndrome Pearson na verdade têm uma mistura de mtDNA deletado e normal, um exemplo de heteroplasmia mitocondrial. Indivíduos que são homoplásmicos para o mtDNA deletado na síndrome Pearson nunca foram observados, provavelmente porque morrem bem no início do desenvolvimento.

#### ■ PONTO IMPORTANTE

■ A neuropatia óptica hereditária Leber e a síndrome medula-pâncreas Pearson são causadas por mutações no mtDNA.

## A GENÉTICA MOLECULAR DOS CLOROPLASTOS

Os cloroplastos contêm moléculas de DNA com uma variedade de genes.

Os cloroplastos são formas especializadas de uma classe geral de organelas vegetais chamadas de **plastídeos**. Botânicos distinguem vários tipos de plastídeos, incluindo cromoplastos (plastídeos contendo pigmentos), amiloplastos (plastídeos contendo amido) e elaioplastos (plastídeos contendo óleo ou lipídeos). Todos os três tipos parecem desenvolver-se de pequenas organelas delimitadas por membrana chamadas de pró-plastídeos e, em uma determinada espécie de planta, todos parecem conter o mesmo DNA. Esse DNA geralmente é referido como **DNA de cloroplasto**, abreviadamente **cpDNA**.

### DNA de Cloroplastos

Em plantas superiores, o cpDNA tipicamente varia de 120 a 160 kb de tamanho e em algas de 85 a 292 kb. Em algumas espécies de algas verdes do gênero *Acetabularia*, o cpDNA é muito grande, com cerca de 2.000 kb. Entre as espécies de plantas cujo DNA de cloroplasto foi pelo menos parcialmente caracterizado, o cpDNA parece ser organizado como uma molécula circular fechada. Entretanto, em algumas es-

pécies, sobretudo aquelas com grandes cpDNA, um a dia, a linear não pode ser excluído.

O número de moléculas de cpDNA em uma célula depende de dois fatores: o número de cloroplastos e o número de moléculas de cpDNA dentro de cada cloroplasto. Por exemplo, na alga unicelular *Chlamydomonas reinhardtii*, só há um cloroplasto por célula, e ele contém cerca de 100 cópias de cpDNA. Em *Euglena gracilis*, outro organismo unicelular, existem cerca de 15 cloroplastos por célula, e cada um contém cerca de 40 cópias de cpDNA.

Todas as moléculas de cpDNA levam basicamente o mesmo conjunto de genes, mas em espécies diferentes esses genes são dispostos de modos diferentes. O conjunto gênico básico inclui genes para RNA ribossômicos, RNA transportadores, algumas proteínas ribossômicas, vários componentes polipeptídicos dos fotossistemas que estão envolvidos em captura de energia solar, a subunidade cataliticamente ativa da enzima ribulose 1,5-bisfosfato-carboxilase e quatro subunidades de uma RNA-polimerase específica de cloroplastos. Algumas moléculas de cpDNA foram totalmente seqüenciadas, incluindo a da hepática, *Marchantia polymorpha* (Fig. 19.15) e outra da planta do tabaco, *Nicotiana tabacum*. O cpDNA de tabaco é maior (155.844 pb) e provavelmente contém cerca de 150 genes. A melhor estimativa para o número de genes no cpDNA de hepática (121.024 pb) é 136. A maioria dos cpDNA tem um par de grandes repetições invertidas que contêm genes para RNA ribossômicos. Essas repetições variam entre 10 e 76 kb de tamanho e estão variadamente situadas em moléculas diferentes de cpDNA.

### Biogênese de Cloroplastos

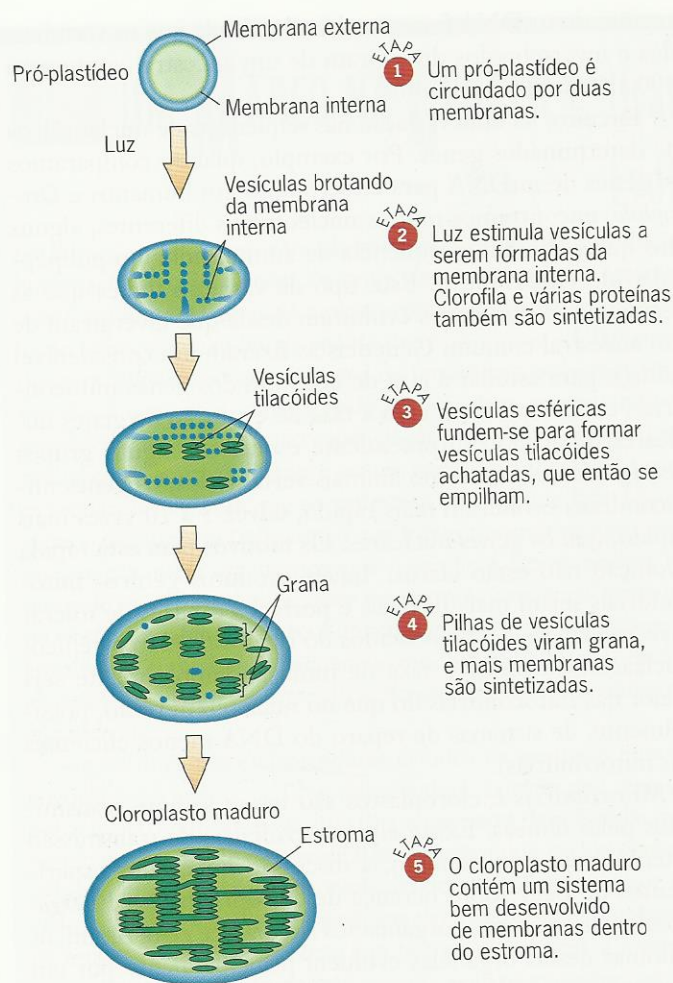
Como já foi mencionado, todos os plastídeos desenvolvem-se de pró-plastídeos. No caso de cloroplastos, este desenvolvimento é estimulado por luz e envolve a transcrição de muitos genes, incluindo alguns situados no núcleo. A iluminação ativa uma profunda mudança em cada pró-plastídeo. A organela aumenta de tamanho, e o interior de suas duas membranas circundantes se expande para brotar de vesículas que se distribuem em pilhas formando estruturas chamadas grana (Fig. 19.16). Todas as proteínas e pigmentos de clorofila necessários para a fotossíntese são feitos e direcionados para seus locais apropriados dentro do cloroplasto emergente. Algumas destas moléculas são posicionadas nas membranas tilacóides que constituem as grana e outras estão situadas no estroma, que é o espaço protoplasmático que circunda a grana.

A formação de cloroplastos funcionais é um processo chamado de **biogênese**. Só alguns detalhes são conhecidos. Entretanto, está claro que a luz tem um papel importante. Por exemplo, um gene nuclear que codifica a pequena subunidade de ribulose 1,5-bisfosfato-carboxilase é vigorosamente transcrito quando é fornecida luz. Uma classe especial de proteínas pigmentadas chamadas **fitocromos** parece mediar esta e outras respostas à luz. Absorvendo energia da luz, tais









**Fig. 19.16 ■ Biogênese de cloroplastos.** Um cloroplasto maduro contendo pilhas de membranas tilacóides (grana) dentro de sua matriz protoplasmática (estroma) desenvolve-se de um pró-plastídeo após exposição a luz.

sintética foi localizada em membranas dentro das células, e não organizada dentro de organelas subcelulares especiais. A fotossíntese gerou quantidades substanciais de  $O_2$ , e, à medida que este gás se acumulou na atmosfera, o metabolismo aeróbico evoluiu, provavelmente há 1 a 2 bilhões de anos. Os primeiros organismos aeróbicos também eram procarióticos.

Os organismos eucarióticos evoluíram há 1 a 1,5 bilhão de anos. Os primeiros eucariontes conhecidos eram algas verdes filamentosas, com núcleos e uma organização intracelular elaborada, incluindo organelas subcelulares. Estes organismos foram capazes de fotossíntese e metabolismo aeróbico e provavelmente continham tanto cloroplastos quanto mitocôndrias. De onde vieram estas organelas citoplasmáticas?

Nos anos 1970, a cientista americana Lynn Margulis questionou que tanto mitocôndrias quanto cloroplastos já foram organismos de vida livre que se tornaram incor-

porados a células eucarióticas primitivas. Hoje em dia, a hipótese de que cloroplastos e mitocôndrias se originaram como organismos separados é amplamente aceita. Tais organismos evidentemente estabeleceram uma relação simbiótica com as células eucarióticas primitivas que as internalizaram.

## Organelas Eucarióticas como Endossimbiontes

**Simbiose** é uma condição na qual dois ou mais organismos vivem juntos em íntima associação para seu mútuo benefício. Os líquens são um exemplo clássico. Estes organismos amplamente distribuídos são compostos de uma alga e um fungo que vivem juntos em uma associação íntima mutuamente benéfica. A **endossimbiose** é um caso especial no qual um dos parceiros simbióticos vive dentro do outro. Por exemplo, células de *Chlorella*, uma alga unicelular, às vezes vive dentro do protozoário unicelular *Paramecium bursaria*. A clorofila nestas algas endossimbióticas faz com que o *Paramecium* pareça verde.

Evidências de estudos bioquímicos, genéticos e moleculares sugerem que mitocôndrias e cloroplastos se originaram como endossimbiontes bacterianos. Mitocôndrias e cloroplastos são mais ou menos do mesmo tamanho de algumas bactérias, e ambos contêm moléculas circulares de DNA como cromossomos bacterianos circulares. Tanto mitocôndrias quanto cloroplastos contêm seus próprios ribossomos distintos, e ambos utilizam uma RNA-polimerase distinta. Em cloroplastos, essa RNA-polimerase é similar à encontrada em *E. coli*. Estudos de seqüenciamento do DNA demonstraram que alguns genes nos DNA mitocondrial e de cloroplasto são similares a genes bacterianos. Por exemplo, os genes mitocondriais para RNA ribossômico são mais proximamente relacionados aos genes de RNA ribossômico em um grupo de bactérias fotossintéticas púrpura e os genes de cloroplastos para proteínas ribossômicas parecem estar relacionados a genes bacterianos para estas proteínas. Todos estes fatos sugerem fortemente que mitocôndrias e cloroplastos se originaram quando bactérias foram incorporadas a células eucarióticas primitivas, provavelmente há mais de um bilhão de anos. As bactérias endossimbióticas beneficiaram-se de um ambiente biológico estável que forneceu alimento e capacidade de se reproduzir, e as células eucarióticas beneficiaram-se das fontes energéticas oxidativas e fotossintéticas das bactérias endossimbióticas.

Desde a formação destas relações, cada um dos parceiros simbióticos sofreu mudanças significativas, e ocorreu um considerável embaralhamento genético entre os DNA mitocondrial, de cloroplasto e nuclear. Por exemplo, o mtDNA em algumas plantas superiores contém um gene de RNA transportador de cpDNA e em outras espécies de plantas o mtDNA contém parte do gene do cloroplasto para a subunidade maior de ribulose 1,5-bisfosfato-carboxilase. Na maioria das espécies de plantas, a subunidade menor de ribulose 1,5-bisfosfato-carboxilase é codificada



por um gene nuclear, mas em algumas espécies de algas ela é codificada por um gene de cloroplasto. Estas diferenças indicam que alguns dos genes originalmente presentes no DNA das bactérias endossimbióticas foram transferidos para outros compartimentos celulares, incluindo muitos que foram transferidos para o núcleo. Foi avaliado que entre 400 e 2.200 dos 25.000 genes projetados no genoma nuclear da planta *Arabidopsis thaliana* são derivados de bactérias, supostamente por transferências dos endossimbiontes que acabaram por tornar-se mitocôndrias e cloroplastos nestas espécies. Em virtude desse embaralhamento gênico, nem mitocôndrias nem cloroplastos são capazes de se manter sem materiais especificados pelo núcleo. A maioria dos materiais necessários para montagem e funcionamento destas organelas são codificados por genes nucleares. Mesmo componentes importantes dos sistemas genéticos de mitocôndrias e cloroplastos são derivados de produtos gênicos nucleares. Eles incluem as DNA- e RNA-polimerases, alguns RNA transportadores e muitas, ou a totalidade, das proteínas ribossômicas. Com o período evolutivo, os parceiros simbióticos mudaram tantos genes que hoje são completamente interdependentes.

## A Evolução de Mitocôndrias e Cloroplastos

A variação entre os DNA de mitocôndrias e cloroplastos indica que os genomas destas organelas sofreram mudanças significativas. Primeiro, há variação no conteúdo gênico. Por exemplo, o mtDNA do protozoário flagelado *Reclinomonas americana* tem 97 genes, 62 dos quais codificam proteínas. Os genes que são encontrados em mtDNA de protistas, animais e fungos são um subgrupo de genes mitocondriais de *Reclinomonas*. Em plantas, o repertório de genes mitocondriais foi aumentado pela aquisição de um grande número de matrizes de leitura não atribuídas, seqüências não codificantes e íntrons. Se o mtDNA de *Reclinomonas* for visto como "primitivo" ou uma condição ancestral, então moléculas de mtDNA de protistas, animais e fungos representam estados "derivados" que são caracterizados por ampla perda de genes. Moléculas de mtDNA de plantas mais provavelmente têm a condição ancestral, pois possuem muitos dos genes encontrados no mtDNA de *Reclinomonas*, mas também têm seqüências adicionais. A evolução do DNA mitocondrial portanto foi caracterizada por mudanças significativas no conteúdo de genes. Mudanças no conteúdo de genes também ocorreram durante a evolução do DNA de cloroplastos.

Segundo, há uma variação estrutural, ou variação na ordem dos genes dentro do DNA. Por exemplo, mtDNA humano e de *Drosophila* contêm o mesmo conjunto de genes, mas estes genes estão dispostos de modo diferente em seus respectivos mtDNA. As disposições diferentes sugerem que

regiões do mtDNA foram realocadas desde que os vertebrados e invertebrados divergiram de um ancestral comum há 400 a 600 milhões de anos.

Terceiro, há uma variação nas seqüências de nucleotídeos de determinados genes. Por exemplo, quando comparamos os genes de mtDNA para citocromo b em humanos e *Drosophila*, encontramos muitos nucleotídeos diferentes, alguns dos quais alteram a seqüência de aminoácidos no polipeptídeo do citocromo b. Este tipo de variação indica que as seqüências destes genes evoluíram desde que divergiram de um ancestral comum. Geneticistas fizeram um considerável esforço para estudar a taxa de evolução dos genes mitocondriais e a compararam com a taxa de evolução de genes nucleares. O achado surpreendente é que, em alguns grupos de organismos tais como animais vertebrados, os genes mitocondriais evoluíram mais rápido, talvez 5 a 10 vezes mais rápido, que os genes nucleares. Os motivos para esta rápida evolução não estão claros. Talvez produtos gênicos mitocondriais sejam mais flexíveis e portanto capazes de tolerar mais mudanças de aminoácidos do que os produtos gênicos nucleares, ou talvez a taxa de mutação simplesmente seja maior nas mitocôndrias do que no núcleo (em razão, possivelmente, de sistemas de reparo do DNA menos eficientes nas mitocôndrias).

Mitocôndrias e cloroplastos são quase sempre transmitidos pelas fêmeas. Existe alguma evidência de transmissão paterna (veja, por exemplo, a discussão inicial dos experimentos de Baur sobre herança de cloroplastos em *Pelargonium*), mas em muitos organismos isto parece raro. Assim, os genomas destas organelas evoluem principalmente por um processo assexuado. Entretanto, como aparentemente não ocorre recombinação entre genomas de organelas dentro de determinadas células, variantes genéticas em duas moléculas diferentes podem, em princípio, ser combinadas. O significado evolucionário desta possibilidade ainda não foi bem avaliado.

O DNA de organelas, especialmente o das mitocôndrias, foi muito útil no estudo da evolução de muitos grupos de organismos. Um Marco na Genética: Usando DNA Mitocondrial para Estudar a Evolução Humana descreve como a análise do mtDNA ajudou a elucidar a evolução de nossa própria espécie.

### PONTOS IMPORTANTES

- Mitocôndrias e cloroplastos parecem ter-se originado como bactérias que foram incorporadas a células eucarióticas há cerca de um bilhão de anos.
- À medida que bactérias endossimbióticas evoluíram em organelas, muitos de seus genes foram perdidos; outros genes foram transferidos para o genoma da célula hospedeira ou para o genoma de outra bactéria endossimbiótica.
- O DNA de organelas, especialmente o mtDNA, tem sido muito útil no estudo da evolução de grupos diferentes de organismos.





## UM MARCO NA GENÉTICA: Usando DNA Mitochondrial para Estudar a Evolução Humana

Em biologia, poucos assuntos são mais fascinantes que a evolução humana. Quem somos nós? De onde viemos? Para onde estamos indo? Antes do advento da biologia molecular, o estudo da evolução humana dependeu da análise de fósseis raros – fragmentos de ossos, alguns dentes, uma arma ou ferramenta ocasional. Hoje em dia, a evolução humana pode ser estudada comparando-se seqüências de DNA. Cada seqüência de DNA descende de uma seqüência que estava presente em um organismo ancestral. Assim, as seqüências de DNA que encontramos hoje são, de fato, fósseis vivos – registros de DNA antigo que foram transmitidos por muitas gerações para organismos que vivem hoje. Como podem ter ocorrido mutações durante este tempo, uma seqüência moderna de DNA provavelmente não é uma réplica exata de sua ancestral. Entretanto, comparando seqüências modernas de DNA, podemos às vezes reconstruir características do processo evolucionário que as produziu.

Alguns dos mais esclarecedores estudos da evolução humana envolveram a análise de DNA mitocondrial. Existem dois motivos pelos quais o mtDNA é tão útil: (1) grande parte dele evoluiu mais rápido do que o DNA nuclear e (2) ele é transmitido exclusivamente pelas mulheres. A rapidez da evolução do mtDNA permite que um cientista detecte mudanças genéticas significativas em um período de tempo relativamente curto (em termos evolutivos), e a estrita transmissão materna do mtDNA permite a um pesquisador rastrear seqüências modernas de DNA até uma mulher ancestral comum.

Estudos pioneiros do mtDNA humano foram feitos nos anos 1980 por Allan Wilson, Rebecca Cann, Mark Stoneking e seus colaboradores. Estes estudos estabeleceram que existe relativamente pouca variação no mtDNA de diferentes populações humanas e que a maior variação é encontrada no mtDNA de populações da África. Em vista da taxa conhecida de evolução do mtDNA, essas descobertas sugerem que seres humanos modernos se originaram bem recentemente, talvez dentro dos últimos 200.000 anos, e provavelmente na África. Embora tais conclusões fossem inicialmente controversas, trabalhos posteriores as reforçaram.<sup>1</sup> O laboratório de Wilson coletou amostras de mtDNA de mais de 200 pessoas representando muitos grupos raciais e étnicos diferentes. As seqüências de mtDNA nesta coleção foram determinadas bioquimicamente e então analisadas por um programa de computação que dispõe as seqüências em uma árvore filogenética, ou evolutiva. A conclusão de Wilson foi surpreendente. O mtDNA em todos os grupos modernos de seres humanos descende de uma molécula de mtDNA que existiu em uma única mulher que viveu na África há cerca de 200.000 anos. Aplicando uma metáfora bíblica, a imprensa popular apelidou esta mulher de “Eva mitocondrial”.

Enfocando a evolução do mtDNA, o laboratório de Wilson rastreou a ancestralidade humana de volta a um ponto em que as linhagens maternas de todas as seqüências modernas de mtDNA convergem para um único ancestral comum, a mãe mitocondrial de todos nós. Entretanto, tais pesquisadores não disseram que apenas uma mulher deu origem a todos os seres humanos modernos. A massa de DNA nuclear humano, que é igualmente herdada de homens e mulheres e que varia entre os membros de uma população reprodutiva, não pode ser relacionada a um único indivíduo.

O trabalho de Wilson e seus colegas defende fortemente que todos os humanos modernos evoluíram de indivíduos que viveram na África há menos de 200.000 anos, e possivelmente tão recentemente quanto 120.000 anos. Os migrantes dessa população africana original supostamente fundaram as primeiras populações humanas da Europa e da Ásia, que, por sua vez, fundaram as primeiras populações da Austrália, da Oceania e das Américas. Este cenário evolutivo foi chamado de hipótese “Fora da África”. Uma outra hipótese propõe que os humanos evoluíram simultaneamente em muitas regiões do mundo a partir de grupos que há muito se haviam estabelecido nessas regiões, talvez por muitas centenas de milhares de anos, e que estes grupos provavelmente se reproduziram com outras populações arcaicas tais como os Neandertais da Europa e do oeste da Ásia.

Os Neandertais sempre foram um grupo enigmático para estudantes de evolução humana. Restos fósseis indicam que eles eram bem diferentes dos humanos modernos; ossos mais espessos, maior musculatura e proporções corpóreas diferentes os distinguiam claramente. Os Neandertais seriam ancestrais dos humanos modernos? Eles se reproduziram com populações que ao final produziram os humanos modernos ou eles eram espécies separadas e distintas?

Em 1997, Matthias Krings, Anne Stone, Ralf Schmitz, Heike Krainitzki, Mark Stoneking e Svante Pääbo publicaram a seqüência de 379 pares de bases do mtDNA extraído de um osso do braço de um Neandertal fossilizado.<sup>2</sup> Este fóssil particular, descoberto em 1856 perto de Dusseldorf, Alemanha, foi objeto de muitos estudos intensivos. Após longas negociações, os responsáveis pelo fóssil deram permissão a Krings e colaboradores para remover um pedaço de osso de 3,5 g do úmero direito. Pequenos fragmentos deste pedaço foram pulverizados, e os restos de DNA dentro deles foram cuidadosamente extraídos. Devido à idade do fóssil (entre 30.000 e 100.000 anos), esperava-se que a maior parte do DNA estivesse degradada. Entretanto, como o mtDNA é muito mais abundante do que qualquer seqüência particular de DNA nuclear, Krings e colaboradores esperavam que parte dele tivesse sobrevi-

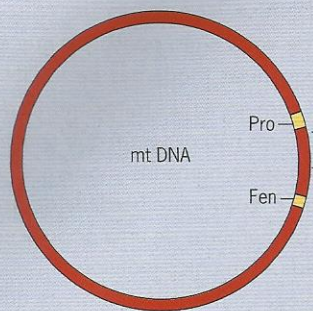
<sup>1</sup>Wilson, A. C., e R. L. Cann. 1992. The recent African genesis of humans. *Sci. Amer.*, 266(4): 68-73.

<sup>2</sup>Krings, M. A. Stone, R. W. Schmitz, H. Krainitzki, M. Stoneking e S. Pääbo. 1997. Neanderthal DNA sequences and the origin of modern humans. *Cell* 90: 19-30.





## UM MARCO NA GENÉTICA (continuação)



GTTCTTTTCATGGGGAGCAGATTGTTGGGTAC  
 CACCCAAGTATTGACTCACCCATCAACAAC  
 CGCTATGTATTCGTACATTACTGCTAGCTCT  
 ACCATGAATATTGTACAGTACCATAAATAC  
 TTGACCTACCTGTAGTACATAAAAACCTAAT  
 CCACATCAAAAACCCCCCTCCCATGCTTACA  
 AAGCAAGTACAGCAATCAACCCTCAACTTAG  
 TCACTACATCAACTCAACTCCAAAGCACTCC  
 CTTACACCCACTAGGATACTCAACAAACCTA  
 CCCACCCTTAGACAGTACATAGTACATAAAG  
 CTATTTACCGTACATAGCACATTACAGTCA  
 AATCCCTTCTCGTCCCATGGATGACCCCC  
 CTCAGATAGGGGTCCCTTG

**Fig. 1** ■ Diferenças de nucleotídeos dentro de uma região não codificante com 379 pb do mtDNA de um fóssil Neandertal e a de um ser humano moderno. A região sequenciada fica entre os genes para tRNA de fenilalanina (Fen) e prolina (Pro). Para cada diferença de nucleotídeo (em destaque), o nucleotídeo superior é encontrado em mtDNA de seres humanos modernos e o inferior é encontrado no mtDNA de Neandertal.

vido. Sua primeira etapa foi usar uma técnica chamada reação em cadeia da polimerase (PCR; veja Cap. 15) para amplificar pequenos segmentos de moléculas sobreviventes de mtDNA. A PCR permite que um pesquisador gere milhões de moléculas idênticas de DNA a partir de apenas algumas moléculas por replicação *in vitro* com uma DNA-polimerase bacteriana. A sequência do DNA amplificado pode então ser determinada bioquimicamente.

Em experimentos cuidadosamente controlados, Krings e colaboradores conseguiram amplificar restos de mtDNA extraídos do fóssil. Análise bioquímica deste material amplificado mostrou que o mtDNA do Neandertal difere do mtDNA de humanos modernos em 28 dos 379 nucleotídeos que foram analisados (Fig. 1). O mtDNA isolado de diferentes humanos modernos mostra tipicamente apenas 8 substituições de nucleotídeos nesta região. Assim, o mtDNA do Neandertal é significativamente diferente do de outros humanos modernos. Análise em computador das sequências de DNA sugeriu que as linhagens de mtDNA de humanos e do Neandertal começaram a evoluir separadamente entre 550.000 e 690.000 anos atrás e que os mtDNA de humanos modernos originaram-se entre 120.000 e 150.000 anos atrás, aparentemente na África. Assim, os Neandertais quase certamente não eram ancestrais dos humanos modernos. Eles evoluíram separadamente e, ao final, tornaram-se extintos.

Na seção de discussão de sua publicação, Krings e co-autores concluem que “A sequência de mtDNA dos Neandertais portanto apóia um cenário no qual humanos modernos surgiram recentemente na África como uma espécie distinta e substituíram os Neandertais com pouco ou nenhum entrecruzamento.” Eles também fazem um adendo: “Deve ser enfatizado que as conclusões acima são baseadas em uma única sequência individual; a recuperação e a análise de sequências de mtDNA de amostras adicionais dos

Neandertais são obviamente desejáveis.”<sup>3</sup> Em março de 2000, uma equipe internacional de pesquisadores relatou a sequência de mtDNA de uma segunda amostra Neandertal, esta de uma criança encontrada em uma caverna no norte do Cáucaso.<sup>4</sup> A sequência de 345 pb que foi obtida desta amostra com 29.000 anos difere de humanos modernos em 22 nucleotídeos. Ela difere da sequência do mtDNA do Neandertal europeu em 12 nucleotídeos. Os dois Neandertais são, portanto, mais correlacionados entre si do que com humanos modernos. Considerados juntos, estes estudos de mtDNA antigo sugerem que Neandertais e humanos modernos têm histórias evolutivas separadas – vias paralelas no tempo que divergiram de um ancestral comum há muito tempo. Uma via terminou quando os Neandertais foram extintos; a outra, a nossa, continua.

### QUESTÕES PARA DISCUSSÕES

1. Amostras muito antigas de mtDNA tais como o material dos Neandertais descrito aqui podem ser amplificadas pela reação em cadeia da polimerase e sequenciadas. Que tipos de problemas estão associados a coleta e interpretação deste tipo de informação de sequência de DNA?
2. Diferenças entre sequências de mtDNA podem ser usadas para deduzir “distâncias evolutivas” entre organismos em uma árvore filogenética. Estas distâncias são biologicamente significativas?

<sup>3</sup>ibid, p. 27.

<sup>4</sup>Ovchinnikov, I. V., A. Götherström, G. P. Romanova, V. M. Kharitonov, K. Liden e W. Goodwin. 2000. Molecular analysis of Neanderthal DNA from the northern Caucasus. *Nature* 404: 490-493.



## Exercícios Básicos Ilustram a análise genética básica.

1. Descreva resumidamente dois exemplos de herança não devidos a genes nucleares.

**Resposta:** Um exemplo é a variação de folhas em plantas. As manchas irregulares de tecidos verde e branco se devem a distribuição de cloroplastos durante as divisões celulares que formam a folha. Tipicamente, o tipo de cor da folha presente no tecido materno é preferencialmente transmitido, e às vezes exclusivamente, para a prole. Outro exemplo é o fenótipo de colônia petite em leveduras. O tamanho pequeno de algumas colônias de levedura se deve a sua incapacidade de obter energia eficientemente de açúcares. O defeito metabólico está nas mitocôndrias, que contêm DNA mutante. Alguns tipos de linhagens de leveduras petite transmitem o fenótipo petite para toda sua prole em cruzamentos com linhagens do tipo selvagem; outras nunca transmitem este fenótipo em tais cruzamentos.

2. Descreva resumidamente a organização molecular do DNA mitocondrial humano.

**Resposta:** O mtDNA humano é organizado como uma molécula circular de 17 kb. Ele contém 37 genes, incluindo alguns para rRNA, tRNA e proteínas envolvidas no metabolismo anaeróbico. Nenhum dos genes contém íntrons. Ambos os filamentos do mtDNA humano são transcritos. Alguns genes são codificados pelo filamento L (leve), enquanto outros são codificados pelo filamento H (pesado, *heavy* em inglês).

3. Que tipos de evidências indicam que mitocôndrias e cloroplastos são derivados de bactérias?

**Resposta:** Algumas características da organização de organelas sugerem uma origem bacteriana: tamanho, dobramento das membranas, presença de ribossomos similares aos encontrados em bactérias. Entretanto, a evidência mais convincente da origem bacteriana de mitocôndrias e cloroplastos vem da análise do DNA nestas organelas. Comparações de seqüências indicam que alguns dos genes em genomas de organelas estão relacionados a genes encontrados em genomas bacterianos.

## Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. Uma linhagem mutante  $mt^+$  de *Chlamydomonas* que era resistente ao antibiótico estreptomicina (*stm-r*) e ao herbicida benomil (*ben-r*) foi cruzada com uma linhagem tipo selvagem  $mt^-$  que era sensível a estas duas substâncias (*stm-s* e *ben-s*). Foram analisadas vinte tétrades da prole quanto a tipo reprodutivo e resistência a estreptomicina e benomil. Quando a análise foi completada, as tétrades foram classificadas em três tipos:

| Tipo 1                     | Tipo 2                     | Tipo 3                     |
|----------------------------|----------------------------|----------------------------|
| $mt^{\square} stm-r ben-r$ | $mt^{\square} stm-r ben-s$ | $mt^{\square} stm-r ben-r$ |
| $mt^{\square} stm-r ben-r$ | $mt^{\square} stm-r ben-s$ | $mt^{\square} stm-r ben-s$ |
| $mt^{\square} stm-r ben-s$ | $mt^{\square} stm-r ben-r$ | $mt^{\square} stm-r ben-r$ |
| $mt^{\square} stm-r ben-s$ | $mt^{\square} stm-r ben-r$ | $mt^{\square} stm-r ben-s$ |
| Número 8                   | 9                          | 3                          |

Quais destas características parecem dever-se a um gene não nuclear?

**Resposta:** Resistência/sensibilidade a estreptomicina. Toda a prole do cruzamento é resistente a estreptomicina, como o genitor  $mt^+$ . Esta herança uniparental é diagnóstica de um gene de cloroplasto. As outras duas características, tipo reprodutivo e resistência/sensibilidade a benomil, segregam 1:1 na prole, indicando que são determinadas por genes nucleares. De fato, pelos números dos três tipos de tétrades, podemos deduzir que os loci *mt* e *ben* estão em cromossomos diferentes. Esta dedução ocorre porque o número de tétrades ditipo parental (Tipo 1) é igual ao número de tétrades ditipo não parental (Tipo 2). Veja no Cap. 8 uma explicação mais completa da análise de tétrades quanto a genes nucleares.

2. Derbeneva e col. (2002 *Am. J. Hum. Genet.* 71: 415-421) se-  
qüenciaram o mtDNA obtido de uma amostra de 30 pessoas

aleutas das Commander Islands, as ilhas mais a oeste da cadeia aleutiana. Foram identificados nove tipos distintos de mtDNA, cada um indicado por um número romano. As diferenças entre estes tipos estão resumidas no quadro seguinte usando o tipo VIII como padrão.

| Tipo | Número na Amostra | Nucleotídeos Diferentes do Tipo VIII |
|------|-------------------|--------------------------------------|
| I    | 13                | 8910, 9667                           |
| II   | 4                 | 6554, 8639, 8910, 9667, 16311        |
| III  | 3                 | 8910, 9667, 16519                    |
| IV   | 3                 | 8910, 9667, 11062                    |
| V    | 1                 | 8460, 8910, 9667                     |
| VI   | 1                 | 5081, 8910, 9667                     |
| VII  | 1                 | 8910, 9667, 10695, 11113             |
| VIII | 3                 | padrão                               |
| IX   | 1                 | 16092                                |

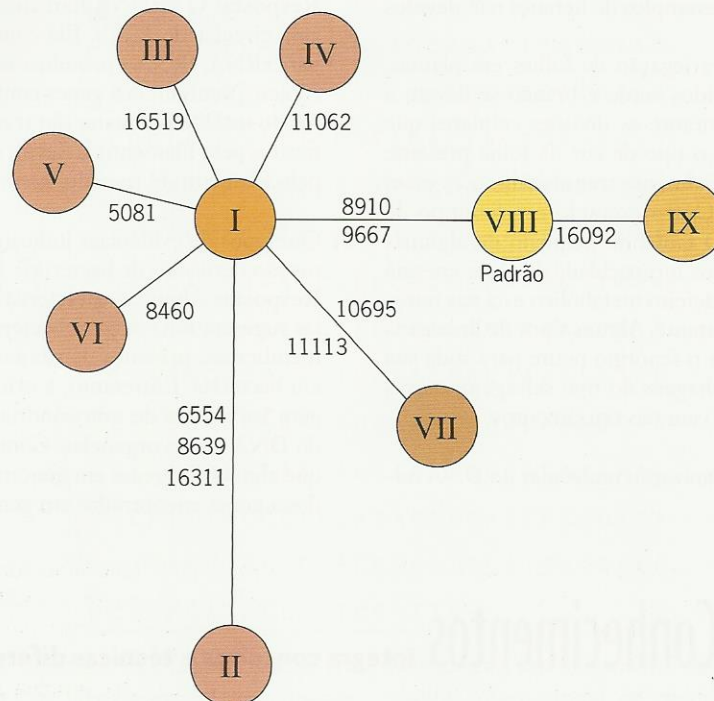
Cada entrada no quadro indica a posição de um nucleotídeo que difere do nucleotídeo encontrado no tipo VIII. Estas diferenças são consistentes entre os vários tipos, isto é, o nucleotídeo diferente na posição 9667 no tipo I é o mesmo que o nucleotídeo diferente na posição 9667 no tipo II. A partir destes dados, construa um diagrama que mostre a correlação filogenética entre os tipos diferentes de mtDNA aleutino.

**Resposta:** O tipo padrão difere do tipo IX em uma posição de nucleotídeo (16092). Ele difere de todos os outros tipos em duas ou mais posições de nucleotídeos – as posições 8910 e 9667 no tipo I, e estas duas posições mais pelo menos uma outra posição em todos os outros tipos. O tipo padrão é portanto uma etapa



mutacional a menos que o tipo IX e é duas etapas mutacionais a menos que o tipo I, mas em um sentido diferente. Todos os outros tipos são uma ou mais etapas mutacionais a menos que

o tipo I. Podemos resumir as correlações entre os tipos em um diagrama filogenético, que, entretanto, não nos diz que mtDNA é ancestral aos outros:



## Questões e Problemas Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 19.1. Quando Correns estudou a herança de variegação em folhas de *Mirabilis*, ele fez cruzamentos entre plantas verdes e variegadas. Por que ele não fez cruzamentos entre plantas verdes e brancas?
- 19.2. A coloração de folhas nas linhagens variegadas é herdada maternamente em *Mirabilis* e biparentalmente em *Pelargonium*. Além disso, em *Pelargonium* o padrão de herança é não mendeliano. O que estes fatos sugerem sobre a herança de cloroplastos nestas espécies?
- 19.3. Explique como plantas tais como *Mirabilis* podem ter setores verde, verde-claro e branco em suas folhas. Se tais setores atingirem a maturidade sexual, que características de cor cada tipo deverá transmitir pelos gametas masculinos e femininos?
- 19.4. Cruzamentos recíprocos com animais ou plantas experimentais às vezes dão resultados diferentes na  $F_1$ . Duas explicações possíveis são (a) herança ligada ao sexo e (b) herança de organelas. Como um pesquisador pode determinar qual explicação é correta?
- 19.5. Oscar Renner fez cruzamentos recíprocos entre dois tipos de primula, *Oenothera bookeri* e *O. muricata*, que se sabe terem a mesma constituição cromossômica. Quando o genitor feminino era *O. bookeri*, as folhas da prole eram amarelas, mas, quando o genitor feminino era *O. muricata*, as folhas da prole eram verdes. Como pode ser explicada esta diferença nos resultados de cruzamentos recíprocos?
- 19.6. Uma linhagem  $mt^+$  de *Chlamydomonas* que era resistente ao antibiótico espectinomicina foi cruzada com uma linhagem sensível  $mt^-$ . Metade da prole era  $mt^+$  e metade  $mt^-$ , mas todas eram resistentes a espectinomicina. Quando as células  $mt^-$  resistentes a espectinomicina foram cruzadas com células  $mt^+$  sensíveis a espectinomicina, toda a prole era sensível a espectinomicina. Explique a herança de resistência/sensibilidade a espectinomicina.
- 19.7. Uma linhagem  $mt^+$  de *Chlamydomonas* que era resistente ao antibiótico espectinomicina foi cruzada com uma linhagem  $mt^-$  que era resistente ao antibiótico estreptomomicina. Sabe-se que os genes para resistência a estes dois antibióticos estão no cpDNA. Antes da reprodução, as células  $mt^+$  foram irradiadas com luz ultravioleta. Cada par cruzado produziu quatro proles haplóides. Em algumas destas tétrades, toda a prole era resistente a espectinomicina e sensível a estreptomomicina. Em outras, toda a prole era resistente a estreptomomicina e sensível a espectinomicina. Em ainda outras, pelo menos uma da prole era resistente a ambos os antibióticos. Explique estes resultados diferentes.



- 19.8. Se o DNA de cloroplastos fosse herdado exclusivamente pelo gameta feminino, seria possível observar recombinação entre genes de cloroplastos?
- 19.9. Uma linhagem mutante de leveduras que produz colônias pequenas quando cultivada em meio rico em glicose foi cruzada com uma linhagem tipo selvagem. Toda a prole era tipo selvagem. Que tipo de mutação foi responsável pelas colônias pequenas?
- 19.10. Uma outra linhagem mutante de leveduras que produziu colônias pequenas foi cruzada com uma linhagem tipo selvagem. Neste cruzamento, entretanto, toda a prole formou colônias pequenas. Que tipo de mutação foi responsável?
- 19.11. Quais são as principais diferenças entre moléculas de mtDNA de plantas e animais?
- 19.12. Se um gene mitocondrial de planta para NADH-redutase fosse inserido em um plasmídeo e então colocado em *E. coli*, ele produziria NADH-redutase cataliticamente ativa? Explique.
- 19.13. Um gene mitocondrial de mamífero foi fundido com um gene de bactéria usando-se técnicas de DNA recombinante. A fusão foi feita de modo que a metade 3' do gene mitocondrial (codificando 278 aminoácidos) é ligada *in frame*\* à metade 5' do gene bacteriano (codificando 200 aminoácidos). Se este gene de fusão for introduzido em bactérias por transformação, ele produzirá um polipeptídeo de fusão com 478 aminoácidos de tamanho? Explique.
- 19.14. Que implicação a *trans*-recomposição de RNA mitocondrial tem para o conceito um gene-um polipeptídeo?
- 19.15. Tanto homens quanto mulheres podem desenvolver a neuropatia óptica hereditária Leber (LHON)? Explique.
- 19.16. Durante o estudo de uma peste de insetos, um geneticista observa quatro formas de uma enzima (A, B, C e D) que podem ser detectadas nos insetos por análise bioquímica. Entretanto, insetos individuais apresentam no máximo duas formas diferentes desta enzima, e, quando duas formas da enzima estão presentes, elas são igualmente abun-

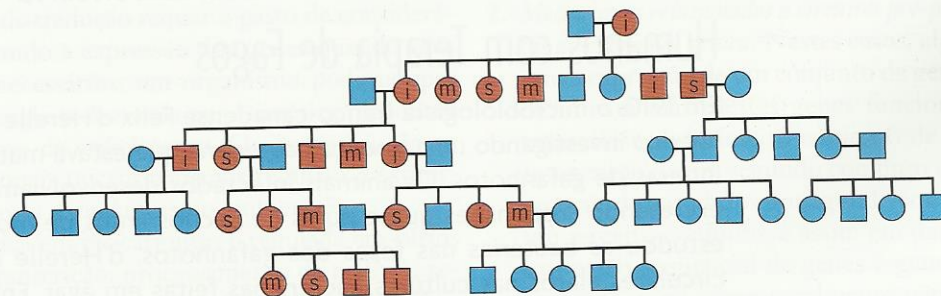
dantes. O geneticista fez uma série de cruzamentos entre insetos que apresentavam formas únicas da enzima. Então a prole F<sub>1</sub> desses cruzamentos foi inter cruzada para produzir a prole F<sub>2</sub>. Pelos dados que foram coletados destes cruzamentos, qual a base genética para as formas diferentes da enzima?

| Cruzamento | Mãe | Pai | F <sub>1</sub> | F <sub>2</sub> (freq. entre parênteses) |
|------------|-----|-----|----------------|-----------------------------------------|
| 1          | A   | B   | A, B           | (1/4) A; (1/2) A, B; (1/4) B            |
| 2          | B   | A   | A, B           | (1/4) A; (1/2) A, B; (1/4) B            |
| 3          | A   | C   | A              | A                                       |
| 4          | C   | A   | C              | C                                       |
| 5          | A   | D   | A, B           | (1/4) A; (1/2) A, B; (1/4) B            |
| 6          | D   | A   | C, D           | (1/4) C; (1/2) C, D; (1/4) D            |
| 7          | C   | D   | C, D           | (1/4) C; (1/2) C, D; (1/4) D            |
| 8          | D   | C   | C, D           | (1/4) C; (1/2) C, D; (1/4) D            |
| 9          | B   | C   | A, B           | (1/4) A; (1/2) A, B; (1/4) B            |
| 10         | C   | B   | C, D           | (1/4) C; (1/2) C, D; (1/4) D            |
| 11         | B   | D   | B              | B                                       |
| 12         | D   | B   | D              | D                                       |

19.17. Em plantas, uma subunidade da enzima ribulose 1,5-bis-fosfato-carboxilase é codificada por um gene nuclear e outra é codificada por um gene de cloroplasto. Ambas as subunidades são necessárias para que a enzima funcione na fotossíntese. Que problema especial a separação dos genes em compartimentos celulares diferentes cria para a síntese desta importante proteína?

19.18. Como o fato de que alguns genomas de organelas possuem íntrons se ajusta à hipótese de que mitocôndrias e cloroplastos são derivados de bactérias endossimbióticas?

19.19. A expressão de cardiomiopatia hereditária em humanos pode ser de branda a grave, com a expressão grave geralmente resultando em morte prematura. O heredograma seguinte mostra a ocorrência e a gravidade de cardiomiopatia em uma família. A gravidade da doença é indicada pelas letras m (branda, *mild*), i (intermediária) e s (grave, *severe*). (a) Que tipo de herança está envolvido em cardiomiopatia? (b) O que poderia explicar a expressividade variável de cardiomiopatia neste heredograma?



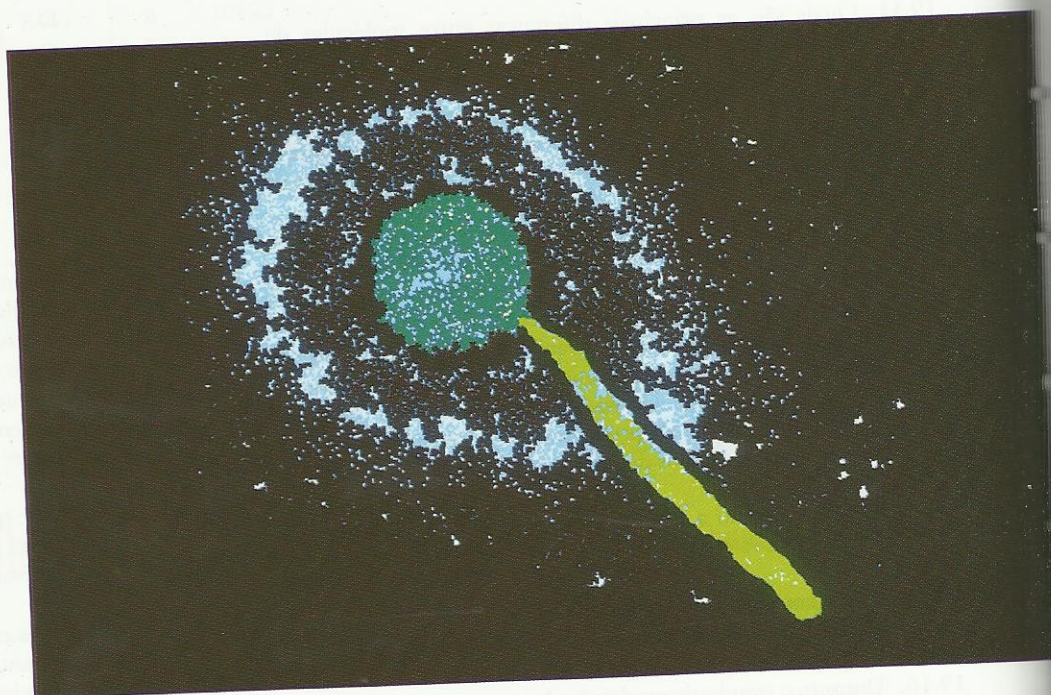
\*N.T.: *In frame* = manutenção de uma matriz de leitura de uma região codificante após uma mutação, ou ligação, de duas moléculas de DNA.



# Regulação da Expressão Gênica em Procariontes e Seus Vírus

## Tópicos

- Expressão Gênica Constitutiva, Indutível e Repressível
- Controle Positivo e Negativo da Expressão Gênica
- Óperons: Regulação Coordenada da Expressão Gênica
- O Óperon de Lactose em *E. coli*: Indução e Repressão Catabólica
- O Óperon de Triptofano em *E. coli*: Repressão e Atenuação
- Bacteriófago Lambda: Lisogenia ou Lise
- Sequências Temporais de Expressão Gênica Durante a Infecção de Fago
- Controle Traducional da Expressão Gênica
- Mecanismos Reguladores Pós-traducionais



Micrografia eletrônica de varredura do bacteriófago *lambda*.

## O Sonho de d'Hérelle de Tratar Disenteria em Seres Humanos com Terapia de Fagos

Em 1910, o microbiologista franco-canadense Felix d'Hérelle encontrava-se no México investigando uma doença bacteriana que estava matando populações inteiras de gafanhotos. Os animais infectados desenvolviam diarreia intensa, excretando suspensões quase puras de bacilos antes de morrer. Quando ele estudou as bactérias nas fezes dos gafanhotos, d'Hérelle observou pontos circulares claros nas culturas bacterianas feitas em ágar. Entretanto, quando examinou o material microscopicamente nos pontos claros, ele não pôde ver nada. Em 1915, d'Hérelle voltou ao Instituto Pasteur em Paris, onde estudou uma epidemia de disenteria bacteriana que estava atacando as unidades do exército estacionadas na França. Ele novamente observou pontos claros nas camadas de bactérias. Além disso, ele demonstrou que o agente causador



de morte das *Shigella*, uma bactéria que causa disenteria em humanos, podia passar por um filtro de porcelana que retinha todas as bactérias conhecidas. Em 1917, d'Hérelle publicou seus resultados e chamou os agentes bactericidas submicroscópicos de bacteriófagos (do grego, "devorador de bactérias"). Mais ou menos ao mesmo tempo, um médico bacteriologista inglês, Frederick W. Twort, descobriu um agente submicroscópico similar que matava micrococos. Infelizmente, a pesquisa de Twort foi logo interrompida por um chamado para servir no Royal Army Medical Corps na Primeira Guerra Mundial.

Enquanto isto, d'Hérelle continuou o estudo dos agentes submicroscópicos que matavam *Shigella*. Ele fez o seguinte relato sobre um de seus experimentos: "... subitamente eu compreendi: o que causava meus pontos claros era de fato um micróbio invisível, um vírus filtrável, mas um vírus parasita de bactérias... Se isto for verdade, a mesma coisa provavelmente ocorreu durante a noite no homem doente, que na véspera estava em grave condição. Em seu intestino, como em meu tubo de ensaio, o bacilo da disenteria dissolvera-se sob a ação de seu parasita. Ele agora deveria estar curado. Eu fui ao hospital. De fato, durante a noite, sua condição havia melhorado muito, e a convalescência estava começando" (d'Hérelle, F. 1949. *The Bacteriophage*. *Science News* 14:44-59).

De fato, d'Hérelle ficou obcecado com sua crença de que doenças humanas causadas por bactérias podiam ser tratadas, talvez mesmo erradicadas, por terapia com bacteriófagos. Infelizmente, logo foi demonstrado que esta forma simples de terapia com bacteriófagos não é eficaz no tratamento de infecções bacterianas, pois muito freqüentemente as bactérias mutam para formas resistentes aos fagos. Entretanto, o trabalho de d'Hérelle preparou o cenário para pesquisas que acabariam produzindo todo um novo campo, a genética microbiana, e esclareceu os mecanismos pelos quais a expressão gênica é regulada. Neste capítulo, examinaremos alguns destes mecanismos.

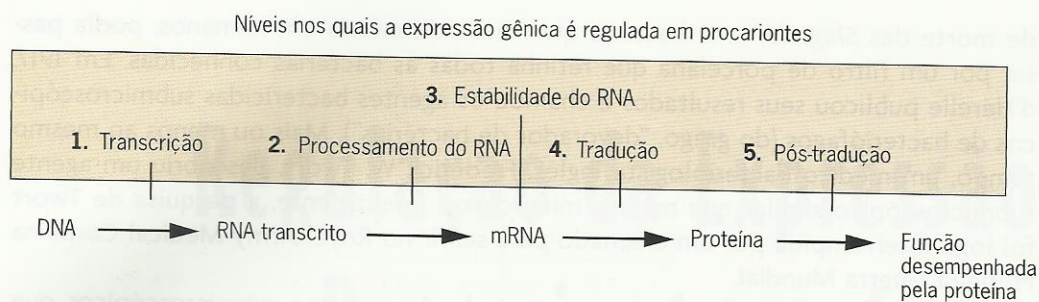
**M**icrorganismos exibem marcantes capacidades em se adaptar a diversas condições experimentais. Essa adaptabilidade depende em parte de sua capacidade de ligar e desligar a expressão de conjuntos específicos de genes em resposta a mudanças no ambiente. A expressão de determinados genes é ligada quando os produtos destes genes são necessários para o crescimento. Sua expressão é desligada quando os produtos gênicos não são mais necessários. A síntese de transcritos gênicos e produtos de tradução requer o gasto de considerável energia. Desligando a expressão dos genes quando seus produtos não são necessários, um organismo pode poupar energia e usar a energia conservada para sintetizar produtos que aumentem a taxa de crescimento. Quais, então, são os mecanismos pelos quais microrganismos regulam a expressão gênica em resposta a mudanças no ambiente?

A expressão gênica em procariontes é regulada em vários níveis diferentes: transcrição, processamento de mRNA, renovação de mRNA, tradução e pós-tradução (Fig. 20.1). Entretanto, os mecanismos reguladores com os maiores efeitos no fenótipo atuam em nível de transcrição.

Com base no que sabemos sobre a regulação da transcrição, os vários mecanismos reguladores parecem enquadrar-se em duas categorias gerais:

1. *Mecanismos que envolvem o rápido ligar e desligar da expressão gênica em resposta a mudanças ambientais.* Mecanismos reguladores deste tipo são importantes em microrganismos devido à freqüente exposição desses organismos a súbitas mudanças no ambiente. Eles fornecem microrganismos com considerável "plasticidade", uma capacidade de ajustar seus processos metabólicos rapidamente para obter o máximo crescimento e reprodução em uma ampla gama de condições ambientais.
2. *Mecanismos relacionados a circuitos pré-programados ou cascatas de expressão gênica.* Nestes casos, alguns eventos dispararam a expressão de um conjunto de genes. O(s) produto(s) de um ou mais destes genes funciona(m) desligando a transcrição do primeiro conjunto de genes ou ligando a transcrição de um segundo conjunto de genes. Então, um ou mais dos produtos do segundo conjunto atuam ligando um terceiro conjunto, e assim em diante. Nestes casos, a expressão sequencial de genes é geneticamente pré-programada, e os genes geralmente não podem ser ligados fora de seqüência. Tais seqüências pré-programadas de expressão gênica estão bem documentadas em procariontes e nos vírus que os atacam. Por exemplo, quando um bacteriófago lítico infecta uma bactéria, os genes virais são expressos em uma seqüência predeterminada, e essa





**Fig. 20.1** ■ Uma via abreviada de expressão gênica, mostrando cinco níveis importantes de regulação em procariontes.

seqüência está diretamente correlacionada com a seqüência temporal de envolvimento do produto gênico na reprodução e na morfogênese do vírus. Na maioria dos exemplos conhecidos de expressão gênica pré-programada, o circuito é cíclico. Por exemplo, durante infecções virais, alguns eventos associados à embalagem do DNA ou do RNA viral na capa de proteínas ativa o programa genético de modo que ocorre a seqüência apropriada de expressão gênica novamente quando uma prole do vírus infecta uma nova célula hospedeira.

## EXPRESSÃO GÊNICA CONSTITUTIVA, INDUTÍVEL E REPRESSÍVEL

Genes que especificam componentes celulares que desempenham funções de manutenção – por exemplo, RNA ribossômico e proteínas envolvidas na síntese de proteínas – são expressos constitutivamente. Outros genes em geral são expressos apenas quando seus produtos são necessários para o crescimento.

Alguns produtos gênicos, tais como moléculas de tRNA, rRNA, proteínas ribossômicas, subunidades de RNA-polimerase e enzimas que catalisam processos metabólicos geralmente referidos como funções de “manutenção” celular, são componentes essenciais de quase todas as células vivas. Genes que especificam produtos deste tipo estão continuamente sendo expressos na maioria das células. Tais genes são ditos expressos constitutivamente e são designados como **genes constitutivos**.

Outros produtos gênicos são necessários para o crescimento celular apenas em algumas condições ambientais. A síntese constitutiva de tais produtos gênicos seria um desperdício, usando energia que poderia ser utilizada para crescimento mais rápido. A evolução dos mecanismos reguladores que permitem a síntese de tais produtos gênicos apenas quando e onde eles são necessários claramente capacitaria os organismos que possuem estes mecanismos reguladores com uma vantagem seletiva em relação aos organismos que não os têm. Isto sem dúvida explica por que organismos atualmente existentes, incluindo bactérias e vírus, exibem mecanismos altamente eficientes para o controle da expressão gênica.

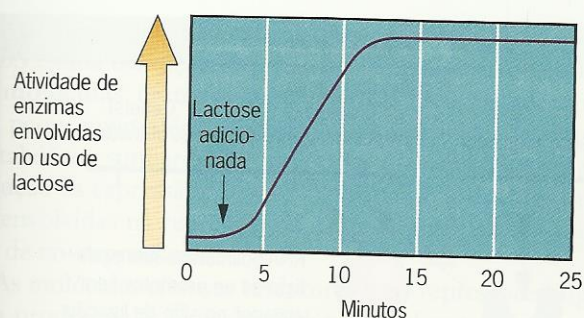
*Escherichia coli* e a maioria das outras bactérias são capazes de crescer usando qualquer um dentre vários carboidratos – por exemplo, glicose, sacarose, galactose, arabinose e lactose – como fonte de energia. Se glicose está presente no ambiente, ela será preferencialmente metabolizada pelas *E. coli*. Entretanto, na ausência de glicose, *E. coli* podem crescer muito bem com outros carboidratos. As células que crescem em meio contendo o açúcar lactose, por exemplo, como única fonte de carbono sintetizam duas enzimas,  $\beta$ -galactosidase e  $\beta$ -galactosídeo-permease, que são necessárias para o metabolismo de lactose. A  $\beta$ -galactosídeo-permease bombeia lactose para dentro da bactéria, onde a  $\beta$ -galactosidase a quebra em glicose e galactose. Nenhuma destas enzimas teria qualquer utilização em *E. coli* se não houvesse lactose disponível para elas. A síntese destas duas enzimas requer considerável energia (sob a forma de ATP e GTP; veja Caps. 11 e 12). Assim, *E. coli* desenvolveram um mecanismo regulador pelo qual a síntese destas enzimas de catabolismo de lactose é ativada na presença de lactose e desligada em sua ausência.

Em ambientes naturais (vias intestinais e esgotos), *E. coli* provavelmente encontram ausência de glicose e presença de lactose relativamente rara. Assim, os genes de *E. coli* que codificam as enzimas envolvidas na utilização de lactose são provavelmente desligados na maior parte do tempo. Se células que crescem em um carboidrato que não seja lactose são transferidas para um meio contendo lactose como única fonte de carbono, elas rapidamente começam a produzir as enzimas necessárias para utilização de lactose (Fig. 20.2a). Este processo de ligar a expressão dos genes em resposta a uma substância no ambiente é chamado de **indução**. Genes cuja expressão é regulada deste modo são chamados de **genes indutíveis**; seus produtos, se enzimas, são chamados de **enzimas indutíveis**.

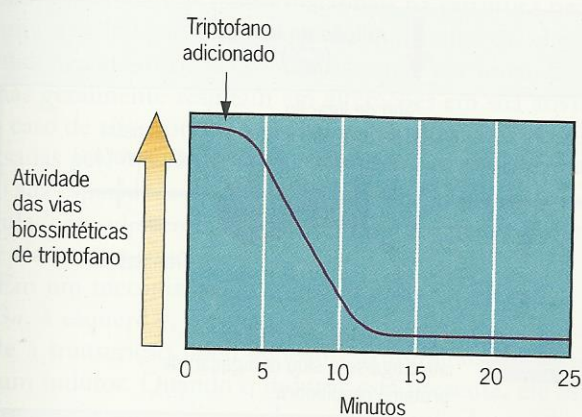
As enzimas que estão envolvidas em **vias catabólicas** (degradativas), como na lactose, galactose ou uso de arabinose, são caracteristicamente indutíveis. Como discutiremos adiante, a indução ocorre no nível da transcrição. A indução altera a taxa de síntese de enzimas, e não a atividade das moléculas enzimáticas existentes. Indução não deve ser confundida com ativação enzimática, que ocorre quando a ligação de uma pequena molécula a uma enzima aumenta a atividade da enzima mas não afeta sua taxa de síntese.

Bactérias podem produzir a maioria das moléculas orgânicas necessárias para crescimento, tais como aminoácidos,





(a) Indução da síntese de enzima.



(b) Repressão da síntese de enzimas.

**Fig. 20.2** ■ (a) Indução de síntese de enzimas necessárias para utilização de lactose como fonte de energia e (b) repressão da síntese de enzimas necessárias para a biossíntese de triptofano, ambas em *E. coli*. Note que ocorrem níveis baixos de síntese de enzimas se os metabólitos estão presentes ou ausentes.

purinas, pirimidinas e vitaminas. Por exemplo, o genoma de *E. coli* contém cinco genes que codificam enzimas catalisadoras de etapas na biossíntese de triptofano. Estes cinco genes devem ser expressos em *E. coli* que crescem em um ambiente desprovido de triptofano com o objetivo de fornecer quantidades adequadas deste aminoácido para a continuidade da síntese de proteínas.

Quando *E. coli* estão presentes em um ambiente contendo triptofano suficiente para apoiar um crescimento ótimo, a síntese contínua das enzimas biossintéticas de triptofano seria um desperdício de energia. Assim, desenvolveu-se um mecanismo regulador em *E. coli* que desliga a síntese das enzimas biossintéticas de triptofano quando há disponibilidade de triptofano externo (Fig. 20.2b). Um gene cuja expressão foi desligada deste modo é dito “reprimido”; o processo é chamado de **repressão**. Quando a expressão deste gene é ligada, ele é dito “desreprimido”; tal resposta é chamada **desrepressão**.

As enzimas que são componentes de **vias anabólicas** (biossintéticas) em geral são repressíveis. A repressão, como a indução, ocorre no nível da transcrição. Repressão não deve ser confundida com inibição por *feedback*, que ocorre

quando o produto de uma via biossintética se liga à atividade da primeira enzima da via e a inibe mas não afeta a síntese da enzima.

### PONTOS IMPORTANTES

- Em procariontes, os genes que especificam funções de manutenção tais como rRNA, tRNA e proteínas ribossômicas são expressos constitutivamente. Outros genes geralmente são expressos apenas quando seus produtos são necessários.
- Genes que codificam enzimas envolvidas em vias catabólicas em geral são expressos apenas na presença de substratos das enzimas; sua expressão é indutível.
- Genes que codificam enzimas envolvidas em vias anabólicas geralmente são desligados na presença do produto final da via; sua expressão é repressível.
- Embora a expressão gênica possa ser regulada em muitos níveis, regulação transcricional é a mais comum.

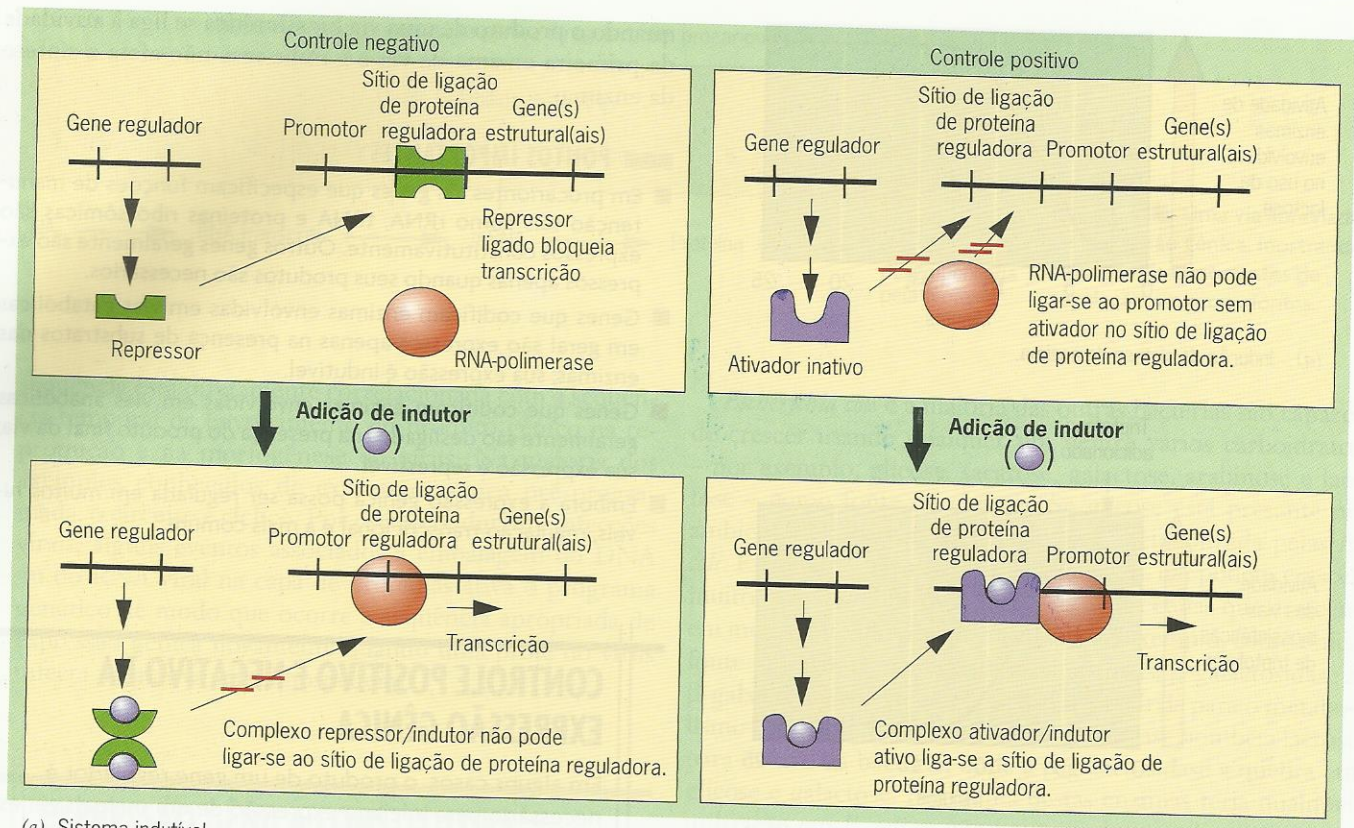
## CONTROLE POSITIVO E NEGATIVO DA EXPRESSÃO GÊNICA

Em alguns casos, o produto de um gene regulador é necessário para iniciar a expressão de um ou mais genes. Em outros casos, o produto de um gene regulador é necessário para desligar a expressão de um ou mais genes.

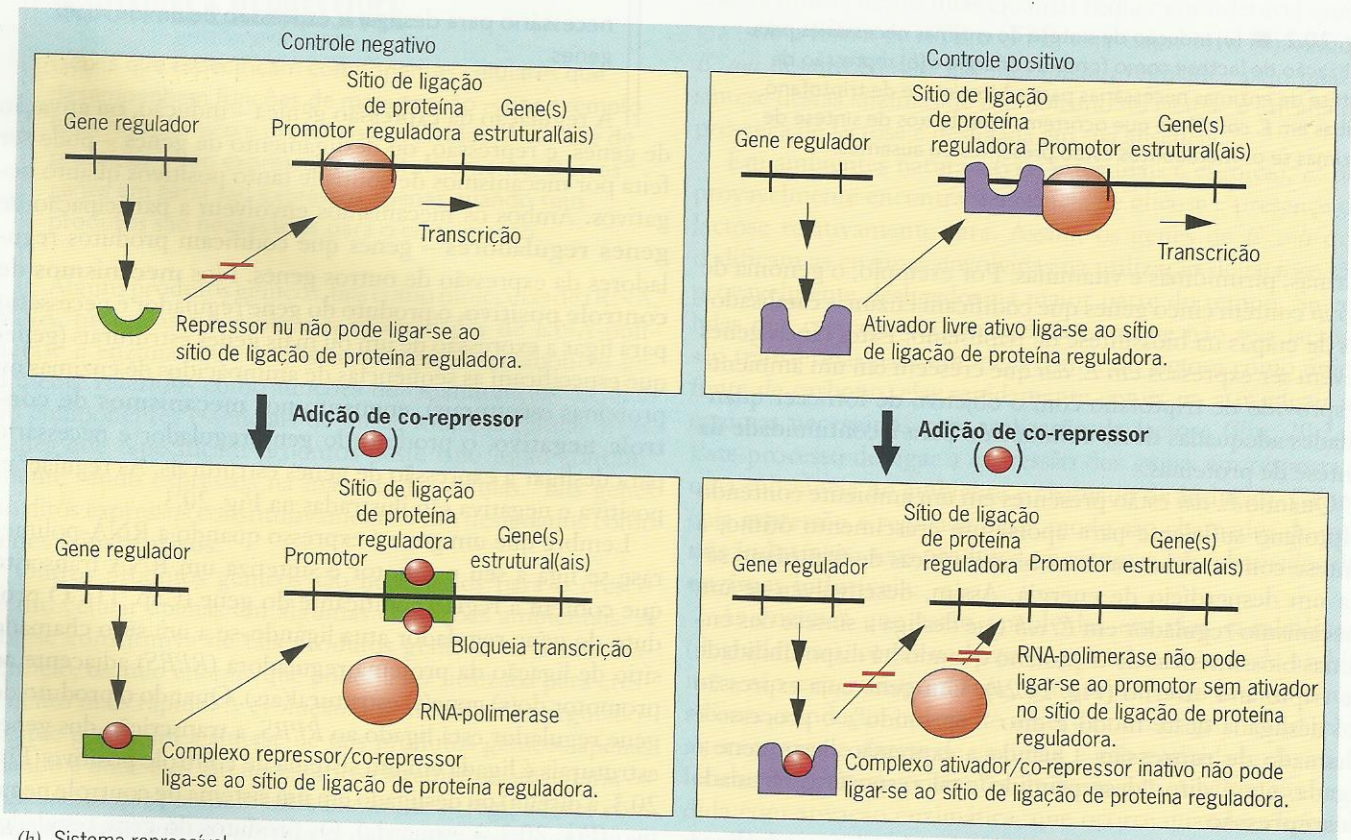
A regulação da expressão gênica – indução, ou ativação de genes, e repressão, ou desligamento de genes – pode ser feita por mecanismos de controle tanto positivos quanto negativos. Ambos os mecanismos envolvem a participação de **genes reguladores** – genes que codificam produtos reguladores da expressão de outros genes. Nos **mecanismos de controle positivo**, o produto do gene regulador é necessário para ligar a expressão de um ou mais genes estruturais (genes que especificam as seqüências de aminoácidos de enzimas ou proteínas estruturais), enquanto nos **mecanismos de controle negativo**, o produto do gene regulador é necessário para desligar a expressão de genes estruturais. As regulações positiva e negativa são ilustradas na Fig. 20.3.

Lembre que um gene é expresso quando a RNA-polimerase se liga a seu promotor e sintetiza um RNA transcrito que contém a região codificante do gene (Cap. 11). O produto do gene regulador atua ligando-se a um sítio chamado sítio de ligação da proteína reguladora (*RPBS*) adjacente ao promotor do(s) gene(s) estrutural(ais). Quando o produto do gene regulador está ligado ao *RPBS*, a transcrição dos genes estruturais é ligada em um sistema de controle positivo (Fig. 20.3, à direita) ou desligada em um sistema de controle negativo (Fig. 20.3, à esquerda). Os produtos do gene regulador são chamados de **ativadores**, pois ativam a expressão gênica, nos sistemas de controle positivo, e **repressores**, porque reprimem a expressão gênica, nos sistemas de controle ne-





(a) Sistema indutível.



(b) Sistema repressível.

**Fig. 20.3 ■** Controle negativo e positivo de expressão gênica indutível e repressível. O produto gênico regulador é necessário para ligar a expressão gênica em circuitos de controle positivo e para desligar a expressão gênica em sistemas de controle negativo.



gativo. Se uma proteína reguladora pode ou não ligar-se ao *RPBS* é uma condição que depende da presença ou ausência de **moléculas efetoras** na célula. Os efetores geralmente são pequenas moléculas tais como aminoácidos, açúcares e metabólitos similares. As moléculas efetoras envolvidas na indução da expressão gênica são chamadas **indutoras**; aquelas envolvidas na repressão da expressão gênica são chamadas de **co-repressoras**.

As moléculas efetoras (indutores e co-repressores) ligam-se a produtos do gene regulador (ativadores e repressores) e causam mudanças nas estruturas tridimensionais destas proteínas. Mudanças conformacionais na estrutura protéica resultantes da ligação de pequenas moléculas são chamadas **transições alostéricas**. Mudanças conformacionais em proteínas geralmente resultam em alterações em sua atividade. No caso de ativadores e repressores, as transições alostéricas causadas pela ligação de moléculas efetoras geralmente alteram sua capacidade de se ligar a sítios de ligação de proteínas reguladoras adjacentes aos promotores dos genes estruturais que elas controlam.

Em um mecanismo de controle indutível negativo (Fig. 20.3a, à esquerda), o repressor livre liga-se ao *RPBS* e impede a transcrição do(s) gene(s) estrutural(ais) na ausência de um indutor. Quando o indutor está presente, ele está ligado ao repressor, e o complexo repressor/indutor não pode ligar-se ao *RPBS*. Sem repressor ligado ao *RPBS*, a RNA-polimerase liga-se ao promotor e transcreve o(s) gene(s) estrutural(ais). Em um mecanismo de controle indutível positivo (Fig. 20.3a, à direita), o ativador não pode ligar-se ao *RPBS* a menos que o indutor esteja presente, e a RNA-polimerase não pode transcrever o(s) gene(s) estrutural(ais) a menos que o complexo ativador/indutor esteja ligado ao *RPBS*. Assim, a transcrição de genes estruturais só é ligada na presença de um indutor.

Em um mecanismo regulador repressível negativo (Fig. 20.3b, à esquerda), a transcrição do(s) gene(s) estrutural(ais) ocorre na ausência de um co-repressor, mas não em sua presença. Quando o complexo repressor/co-repressor está ligado ao *RPBS*, ele impede que a RNA-polimerase transcreva os genes estruturais. Na ausência de co-repressor, o repressor livre pode ligar-se ao *RPBS*. Assim, a RNA-polimerase pode ligar-se ao promotor e transcrever os genes estruturais. Em um mecanismo de controle repressível positivo (Fig. 20.3b, à direita), o produto do gene regulador, o ativador, deve ligar-se ao *RPBS* de modo a que a RNA-polimerase se ligue ao promotor e transcreva o(s) gene(s) estrutural(ais). Quando o co-repressor está presente, ele forma um complexo com a proteína ativadora, e esse complexo ativador/co-repressor é incapaz de se ligar ao *RPBS*. Conseqüentemente, a RNA-polimerase não pode ligar-se ao promotor e transcrever o(s) gene(s) estrutural(ais).

Para compreender os detalhes destes quatro mecanismos de regulação, enfoque as diferenças principais entre eles. (1) O produto do gene regulador, o ativador, participa em ligar a expressão gênica em um mecanismo de controle positivo, enquanto o produto do gene regulador, o repressor, está en-

volvido em desligar a expressão gênica em um mecanismo de controle negativo. (2) Nos dois mecanismos de controle positivo e negativo, se a expressão gênica é indutível ou repressível é uma condição que depende de a proteína reguladora livre ou o complexo proteína reguladora/molécula efetora unir-se ou não ao sítio de ligação da proteína reguladora (*RPBS*).

### PONTOS IMPORTANTES

- A expressão gênica é controlada por mecanismos reguladores positivos e negativos.
- Nos mecanismos de controle positivos, o produto de um gene regulador, um ativador, é necessário para ligar a expressão do(s) gene(s) estrutural(ais).
- Nos mecanismos de controle negativos, o produto do gene regulador, um repressor, é necessário para desligar a expressão do(s) gene(s) estrutural(ais).
- Ativadores e repressores regulam a expressão gênica ligando-se a sítios adjacentes aos promotores de genes estruturais.
- O fato de as proteínas reguladoras poderem ou não ligar-se a seus sítios de ligação depende da presença ou da ausência de pequenas moléculas efetoras que formam complexos com as proteínas reguladoras.
- As moléculas efetoras são chamadas de indutores nos sistemas indutíveis e co-repressores nos sistemas repressíveis.

## ÓPERONS: REGULAÇÃO COORDENADA DA EXPRESSÃO GÊNICA

Em procariontes, genes com funções correlatas geralmente estão presentes em unidades genéticas reguladas coordenadamente chamadas **óperons**.



O **modelo do óperon**, um mecanismo de controle negativo, foi desenvolvido em 1961 por François Jacob e Jacques Monod para explicar a regulação de genes necessários para utilização de lactose em *E. coli*. Discutiremos alguns

dos resultados experimentais que levaram ao desenvolvimento deste modelo em Um Marco na Genética: Jacob, Monod e o Modelo do Óperon. Jacob e Monod propuseram que a transcrição de um conjunto de genes estruturais contíguos é regulada por dois elementos controladores (Fig. 20.4a). Um dos elementos, o gene repressor, codifica um repressor, que (em condições apropriadas) se liga ao segundo elemento, o **operador**. O operador é sempre contíguo aos genes estruturais cuja expressão ele regula. Alguns óperons, incluindo o óperon da lactose discutido na seção seguinte, contêm vários operadores. Entretanto, no momento, consideraremos apenas um só operador para simplificar ao máximo o mecanismo.

A transcrição é iniciada em promotores situados antecedentes (a 5') às regiões codificantes dos genes estruturais.



Quando o repressor está ligado ao operador, ele impede esteiricamente que a RNA-polimerase transcreva os genes estruturais no óperon. As regiões de operador são contíguas com as regiões de promotor; às vezes, operadores e promotores até estão superpostos, compartilhando uma sequência curta de DNA. As regiões do operador em geral estão situadas entre os promotores e os genes estruturais que eles regulam. A unidade contígua completa, que inclui os genes estruturais, o operador e o promotor é chamada **óperon** (Fig. 20.4a).

Se o repressor irá ligar-se ao operador e desligar a transcrição dos genes estruturais em um óperon é uma condição determinada pela presença ou ausência de moléculas efetoras, como discutido na seção anterior. Óperons indutíveis e óperons repressíveis podem ser distinguidos uns dos outros determinando-se se o repressor puro ou o complexo repressor/molécula efetora está ativo em se ligar ao operador.

1. No caso de um óperon indutível, o repressor livre liga-se ao operador, desligando a transcrição (Fig. 20.4b).
2. No caso de um óperon repressível, a situação é revertida. O repressor livre não pode ligar-se ao operador. Apenas o complexo repressor/molécula efetora (co-repressor) é ativo em se ligar ao operador (Fig. 20.4c).

À exceção desta diferença no comportamento de ligação ao operador do repressor livre e do complexo repressor/molécula efetora, óperons indutíveis e repressíveis são idênticos.

Um único mRNA transcrito leva a informação codificante de todo um óperon. Assim, os mRNA de óperons que consistem em mais de um gene estrutural são multigênicos. Por exemplo, o mRNA do óperon de triptofano de *E. coli* contém as sequências codificantes de cinco genes diferentes. Como eles são co-transcritos, todos os genes estruturais em um óperon são coordenadamente expressos. Embora as quantidades molares dos diferentes produtos gênicos não precisem ser iguais (por causa de eficiências diferentes de início da tradução), as quantidades relativas de polipeptídeos diferentes especificados por genes em um óperon geralmente permanecem as mesmas, independente do estado de indução ou repressão. Em alguns casos, o uso diferencial dos sinais de término de transcrição pode alterar as quantidades dos produtos gênicos sintetizados.

### PONTOS IMPORTANTES

- Em bactérias, genes com funções correlatas frequentemente ocorrem em unidades reguladas coordenadamente chamadas óperons.
- Cada óperon contém um conjunto de genes estruturais contíguos, um promotor (o sítio de ligação para a RNA-polimerase) e um operador (o sítio de ligação para uma proteína reguladora chamada repressor).
- Quando um repressor está ligado ao operador, a RNA-polimerase não pode transcrever os genes estruturais no óperon. Quando o operador está livre de repressor, a RNA-polimerase pode transcrever o óperon.

## O ÓPERON DE LACTOSE EM *E. COLI*: INDUÇÃO E REPRESSÃO CATABÓLICA

Os genes estruturais no óperon *lac* são transcritos apenas quando lactose está presente e glicose está ausente.

Jacob e Monod propuseram o modelo do óperon com base em seus estudos do óperon de lactose (*lac*) em *E. coli* (veja Um Marco na Genética: Jacob, Monod e o Modelo do Óperon). O óperon *lac* contém um promotor (*P*), um operador (*O*) e três genes estruturais, *lacZ*, *lacY* e *lacA*, que codificam as enzimas β-galactosidase, β-galactosídeo-permease e β-galactosídeo-transacetilase, respectivamente (Fig. 20.5). A β-galactosídeo-permease “bombeia” lactose para a célula, onde a β-galactosidase a quebra em glicose e galactose (Fig. 20.6). O papel biológico da transacetilase é desconhecido.

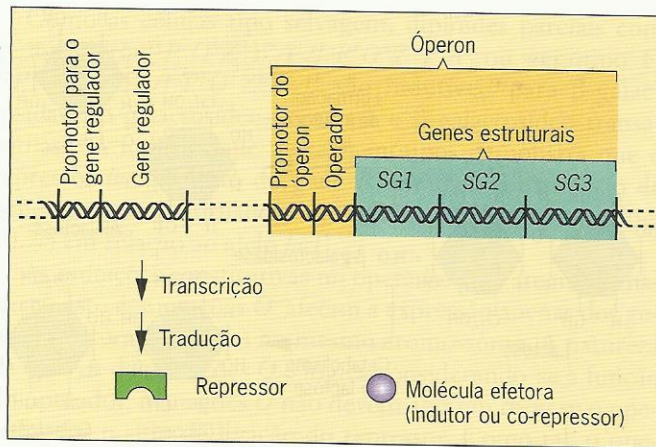
No modelo de Jacob e Monod, o óperon *lac* continha um único operador (agora chamado *O<sub>1</sub>*). Entretanto, dois operadores adicionais (*O<sub>2</sub>* e *O<sub>3</sub>*) foram subsequentemente descobertos. Inicialmente, acreditava-se que *O<sub>2</sub>* e *O<sub>3</sub>* tinham papéis secundários. Então, Benno Müller-Hill e colaboradores demonstraram que a deleção de ambos os operadores “secundários” exercia um grande efeito no nível de transcrição do óperon. Estudos mais recentes mostraram que repressão eficiente do óperon *lac* requer o operador principal (*O<sub>1</sub>*) e pelo menos um dos operadores secundários (*O<sub>2</sub>* ou *O<sub>3</sub>*). Entretanto, primeiro desenvolveremos o modelo de Jacob e Monod do óperon *lac* com um só operador. Depois, ampliaremos o modelo e examinaremos as funções de todos os três operadores em Enfoque Técnico: Interações Proteína-DNA que Controlam a Transcrição do Óperon *lac*.

### Indução

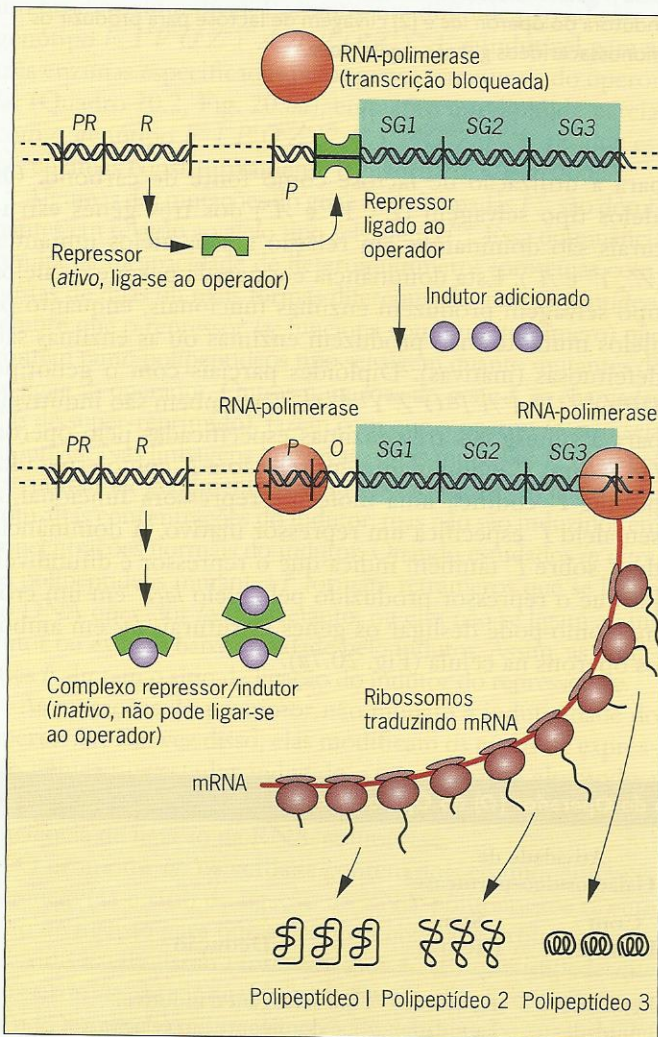
O óperon *lac* é um óperon indutível controlado negativamente; os genes *lacZ*, *lacY* e *lacA* só são expressos na presença de lactose. O gene regulador *lac*, designado como gene *I*, codifica um repressor que tem 360 aminoácidos de tamanho. Entretanto, a forma ativa do repressor *lac* é um tetrâmero contendo quatro cópias do produto do gene *I*. Na ausência do indutor, o repressor liga-se ao operador *lac*, que por sua vez impede que a RNA-polimerase catalise a transcrição dos três genes estruturais (veja a Fig. 20.4b). Algumas moléculas dos produtos gênicos *lacZ*, *lacY* e *lacA* são sintetizadas no estado não induzido, fornecendo baixo nível de atividade enzimática. Essa atividade básica é essencial para indução do óperon *lac* porque o indutor do óperon, a alolactose, é derivado de lactose em uma reação catalisada pela β-galactosidase (Fig. 20.6). Uma vez formada, a alolactose é ligada ao repressor, causando a saída do repressor do operador. Deste modo, a alolactose induz a transcrição dos genes estruturais *lacZ*, *lacY* e *lacA* (veja a Fig. 20.4b).

O gene *lacI*, o operador *lac* e o promotor *lac* foram todos identificados inicialmente mediante procedimentos genéti-

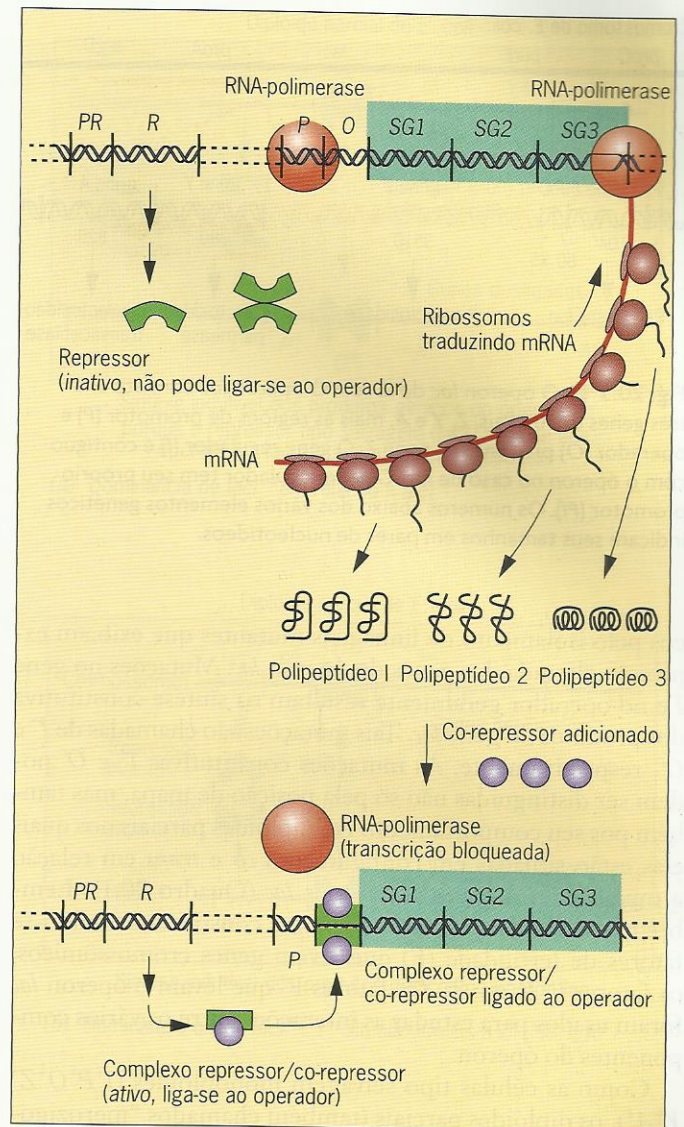




(a) O operon: componentes.



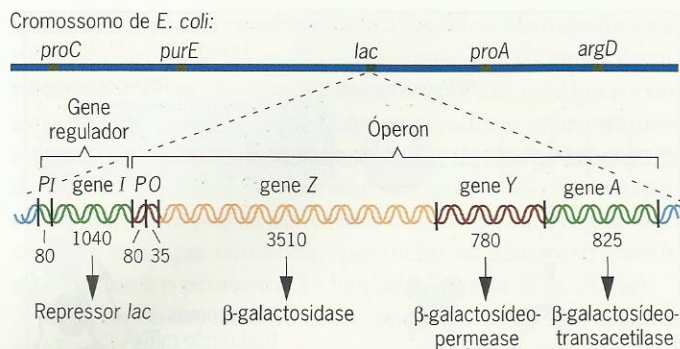
(b) O operon: indução.



(c) O operon: repressão.

**Fig. 20.4** ■ Regulação da expressão gênica pelo mecanismo do operon. (a) Componentes de um operon: um ou mais genes estruturais (são mostrados três, SG1, SG2 e SG3) e as seqüências adjuntas operador (O) e promotor (P). A transcrição do gene regulador (R) é iniciada pela RNA-polimerase, que se liga a seu promotor (PR). Quando o repressor está ligado ao operador, ele impede estericamente que a RNA-polimerase inicie a transcrição dos genes estruturais. A diferença entre um operon indutível (b) e um operon repressível (c) é que o repressor livre se liga ao operador de um operon indutível, enquanto o complexo repressor/molécula efetora se liga ao operador de um operon repressível. Assim, um operon indutível é desligado na ausência da molécula efetora (indutor), e um operon repressível é ligado na ausência da molécula efetora (co-repressor).

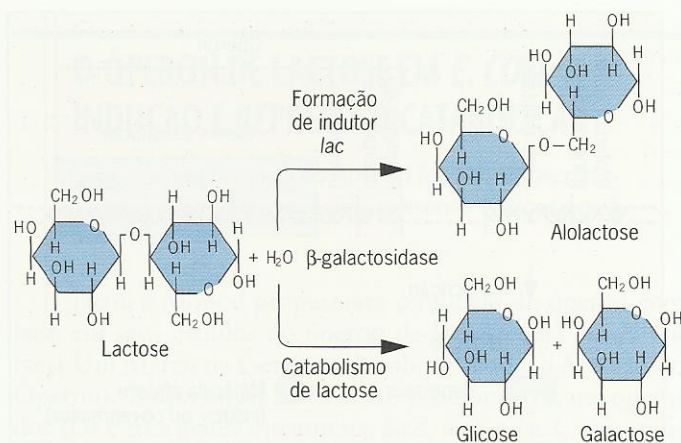




**Fig. 20.5** ■ O operon *lac* de *E. coli*. O operon *lac* consiste em três genes estruturais, Z, Y e A, mais as regiões de promotor (P) e operador (O) próximas ao gene Z. O gene regulador (I) é contíguo com o operon no caso de *lac*. O gene regulador tem seu próprio promotor (PI). Os números abaixo dos vários elementos genéticos indicam seus tamanhos em pares de nucleotídeos.

cos pelo isolamento de linhagens mutantes que exibiam expressão alterada dos genes do operon *lac*. Mutações no gene *I* e no operador geralmente resultam na síntese constitutiva dos produtos do gene *lac*. Tais mutações são chamadas de  $I^-$  e  $O^c$ , respectivamente. As mutações constitutivas  $I^-$  e  $O^c$  podem ser distinguidas não só pela posição de mapa, mas também por seu comportamento em diplóides parciais nos quais elas estão situadas em configurações *cis* e *trans* em relação a mutações nos genes estruturais *lac* (Quadro 20.1). Lembre que diplóides parciais podem ser construídos usando-se fatores de fertilidade (F) que levam genes cromossômicos, os fatores F' (Cap. 8). Os fatores F' que levam o operon *lac* foram usados para estudar as interações entre os vários componentes do operon.

Como as células tipo selvagem monoplóides ( $I^+P^+O^+Z^+Y^+A^+$ ), os diplóides parciais (também chamados “merozigotos”) com genótipo F'  $I^+P^+O^+Z^+A^+/I^-P^+O^+Z^-Y^-A^-$  ou com o genótipo F'  $I^+P^+O^+Z^-Y^-A^-/I^-P^+O^+Z^+Y^+A^+$  são indutíveis



**Fig. 20.6** ■ Duas reações fisiologicamente importantes catalisadas pela β-galactosidase: (1) conversão de lactose em alolactose indutora do operon *lac* e (2) clivagem de lactose para produzir os monossacarídeos glicose e galactose.

para a utilização de lactose como fonte de carbono. Os alelos tipo selvagem ( $Z^+$ ,  $Y^+$  e  $A^+$ ) dos três genes estruturais são dominantes em relação a seus alelos mutantes ( $Z^-$ ,  $Y^-$  e  $A^-$ ). Esta dominância é esperada porque os alelos tipo selvagem produzem enzimas funcionais, enquanto os alelos mutantes não produzem enzimas ou as enzimas são defeituosas (inativas). Diplóides parciais com o genótipo  $I^+P^+O^+Z^+Y^+A^+/I^-P^+O^+Z^-Y^-A^-$  ( $I^+/I^-$ ) também são indutíveis para a síntese das três enzimas especificadas pelo operon *lac*. Assim,  $I^+$  é dominante em relação a  $I^-$  como esperado, porque  $I^+$  codifica uma molécula repressora funcional, e seu alelo  $I^-$  especifica um repressor inativo. A dominância de  $I^+$  sobre  $I^-$  também indica que o repressor é difundível, porque o repressor produzido pelo alelo  $lacI^+$  em um cromossomo pode desligar os genes estruturais *lac* em ambos os operons na célula (Fig. 20.7a).

## QUADRO 20.1

### Efeitos Fenotípicos de Mutações no Gene Repressor (I) e na Região do Operador (O) do Operon lac

| Genótipo                             | Atividade de β-Galactosidase <sup>a</sup> |              | Atividade de β-Galactosídeo-permease <sup>a</sup> |              | Dedução                                                                                                   |
|--------------------------------------|-------------------------------------------|--------------|---------------------------------------------------|--------------|-----------------------------------------------------------------------------------------------------------|
|                                      | Com Lactose                               | Sem Lactose  | Com Lactose                                       | Sem Lactose  |                                                                                                           |
| $I^+P^+O^+Z^+Y^+$                    | 100 unidades                              | 1 unidade    | 100 unidades                                      | 1 unidade    | Tipo selvagem é indutível<br>$Z^+$ é dominante em relação a $Z^-$<br>$Y^+$ é dominante em relação a $Y^-$ |
| $I^+P^+O^+Z^+Y^+/F' I^+P^+O^+Z^-Y^-$ | 100 unidades                              | 1 unidade    | 100 unidades                                      | 1 unidade    |                                                                                                           |
| $I^+P^+O^+Z^+Y^+/F' I^+P^+O^+Z^+Y^+$ | 200 unidades                              | 2 unidades   | 200 unidades                                      | 2 unidades   | Atividade depende de dosagem gênica<br>Mutantes $lacI^-$ são constitutivos                                |
| $I^+P^+O^+Z^+Y^+/F' I^+P^+O^+Z^-Y^-$ | 100 unidades                              | 100 unidades | 100 unidades                                      | 100 unidades |                                                                                                           |
| $I^+P^+O^+Z^+Y^+/F' I^+P^+O^+Z^+Y^+$ | 200 unidades                              | 2 unidades   | 200 unidades                                      | 2 unidades   | $I^+$ é dominante em relação a $I^-$<br>Mutantes $lacO^c$ são constitutivos                               |
| $I^+P^+O^cZ^+Y^+$                    | 100 unidades                              | 100 unidades | 100 unidades                                      | 100 unidades |                                                                                                           |
| $I^+P^+O^cZ^-Y^-/F' I^+P^+O^+Z^+Y^+$ | 100 unidades                              | 100 unidades | 100 unidades                                      | 1 unidade    | $O^c$ e $O^+$ são reguladores de ação <i>cis</i>                                                          |

<sup>a</sup>Os níveis de atividade em bactérias tipo selvagem foram estabelecidos em 100 unidades tanto para β-galactosidase (o produto do gene Z) quanto para β-galactosídeo-permease (o produto do gene Y). O gene A e seu produto β-galactosídeo-transacetilase não são mostrados por motivo de abreviação.



Como as células tipo selvagem, diplóides parciais com genótipo  $F' I^+ P^+ O^+ Z^+ Y^+ A^+ / I^- P^+ O^+ Z^- Y^- A^-$  ou genótipo  $F' I^+ P^+ O^+ Z^- Y^- A^- / I^- P^+ O^+ Z^+ Y^+ A^+$  são indutíveis para  $\beta$ -galactosidase,  $\beta$ -galactosídeo-permease e  $\beta$ -galactosídeo-transacetilase. A indutibilidade destes genótipos demonstra que o repressor *lac* (produto do gene  $I^+$ ) controla a expressão de genes estruturais situados em *cis* (Fig. 20.7b) ou *trans* (Fig. 20.7c) em relação ao alelo  $lacI^+$ .

As mutações constitutivas no operador ( $O^c$ ) atuam apenas em *cis*; isto é, mutações  $O^c$  afetam a expressão apenas dos genes estruturais situados no mesmo cromossomo. A natureza da ação *cis* de mutações  $O$  é lógica considerando-se a função do operador. Mutações  $O$  não devem agir em *trans* se o operador for o sítio de ligação do repressor. Como tal, o operador não codifica nenhum produto, difundível ou não. Um gene regulador deve agir em *trans* apenas se ele especificar um produto difundível. Portanto, um diplóide parcial com o genótipo  $F' I^+ P^+ O^c Z^- Y^- A^- / I^- P^+ O^+ Z^+ Y^+ A^+$  é indutível para as três enzimas especificadas pelos genes estruturais do operon *lac* (Quadro 20.2, Fig. 20.8a), enquanto um diplóide parcial com o genótipo  $F' I^+ P^+ O^c Z^+ Y^+ A^+ / I^- P^+ O^- Z^- Y^- A^-$  produz estas enzimas constitutivamente (Quadro 20.2, Fig. 20.8b).

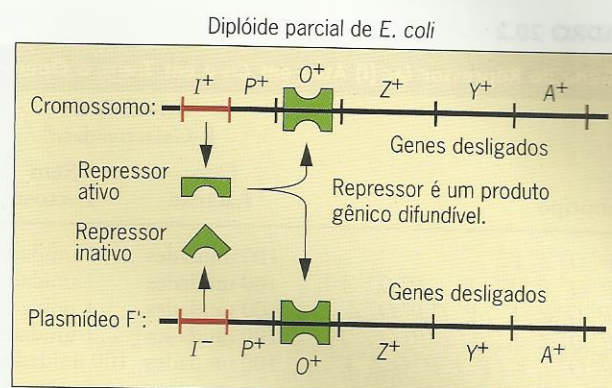
Algumas das mutações do gene *I*, as chamadas  $I^d$ , são dominantes em relação ao alelo tipo selvagem ( $I^+$ ). Essa dominância resulta da incapacidade de heteromultímeros (proteínas compostas de duas ou mais formas diferentes de um polipeptídeo; lembre que o repressor *lac* funciona como um tetrâmero) que contêm tanto polipeptídeos tipo selvagem quanto mutantes de se ligar ao operador. Outras mutações do gene *I*, as chamadas  $I^s$  (de super-reprimido), fazem com que o operon *lac* não seja indutível. Nas linhagens que levam estas mutações  $I^s$ , os genes estruturais *lac* geralmente podem ser induzidos em certo grau com altas concentrações do indutor, mas não são induzidos em concentrações normais do indutor. Quando estudados *in vitro*, os polipeptídeos mutantes  $I^s$  formam tetrâmeros que se ligam ao DNA operador *lac*. Entretanto, ou eles não se ligam ao indutor ou exibem baixa afinidade pelo indutor. Assim, as mutações  $I^s$  alteram o sítio de ligação do indutor do repressor *lac*.

As mutações do promotor não mudam a indutibilidade do operon *lac*. Em vez disso, elas modificam os níveis de expressão gênica nos estados induzido e não induzido alterando a frequência de iniciação da transcrição do operon *lac*, isto é, a eficiência de ligação da RNA-polimerase.

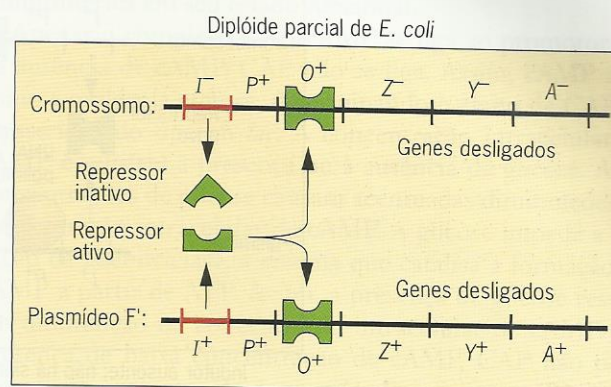
O promotor *lac* na verdade contém dois componentes separados: (1) o sítio de ligação da RNA-polimerase e (2) um sítio de ligação para outra proteína chamada proteína ativadora do catabolismo (abreviada por CAP) que impede que o operon *lac* seja induzido na presença de glicose. Este segundo circuito de controle, que consideraremos em seguida, garante a utilização preferencial de glicose como fonte de energia quando ela está disponível.

## Repressão Catabólica

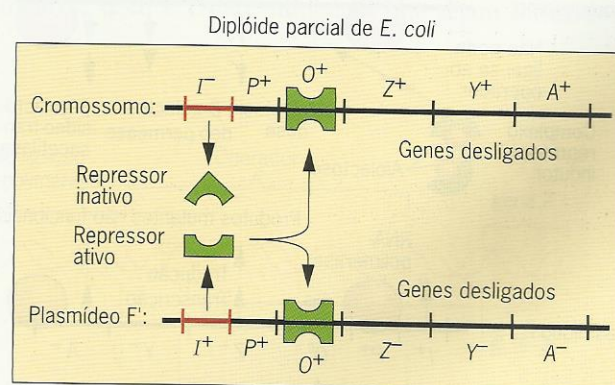
Há muito sabemos que a presença de glicose impede a indução do operon *lac*, bem como outros operons que con-



(a) Dominância de  $lacI^+$  em relação a  $lacI^-$ .



(b) Dominância *cis* de  $lacI^+$ :  $I^+$  situado em *cis* a  $Z^+$ ,  $Y^+$  e  $A^+$ .



(c) Dominância *trans* de  $lacI^+$ :  $I^+$  situado em *trans* a  $Z^+$ ,  $Y^+$  e  $A^+$ .

**Fig. 20.7** ■ O gene  $lacI^+$  é dominante em relação aos alelos  $lacI^-$  (a) e controla os operadores *lac* situados em *cis* (b) ou *trans* (c) em relação a si mesmo. Estes efeitos demonstram que o produto do gene *lac* é difundível. Embora a forma funcional do repressor *lac* seja um tetrâmero, as duas moléculas na parte posterior do tetrâmero não são mostradas por motivo de simplificação.

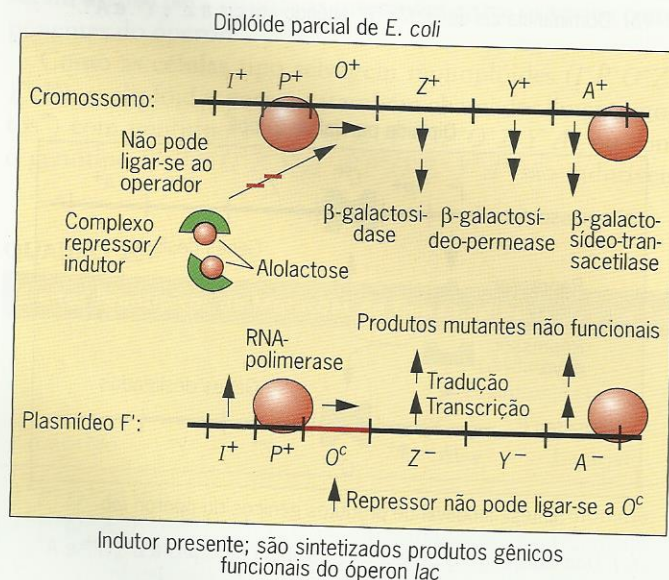
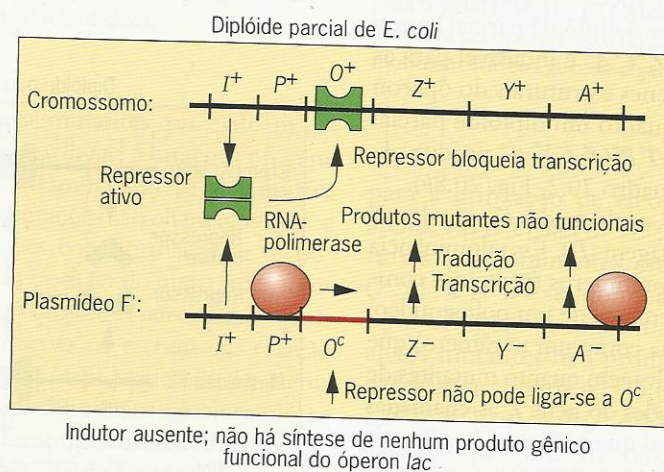


## QUADRO 20.2

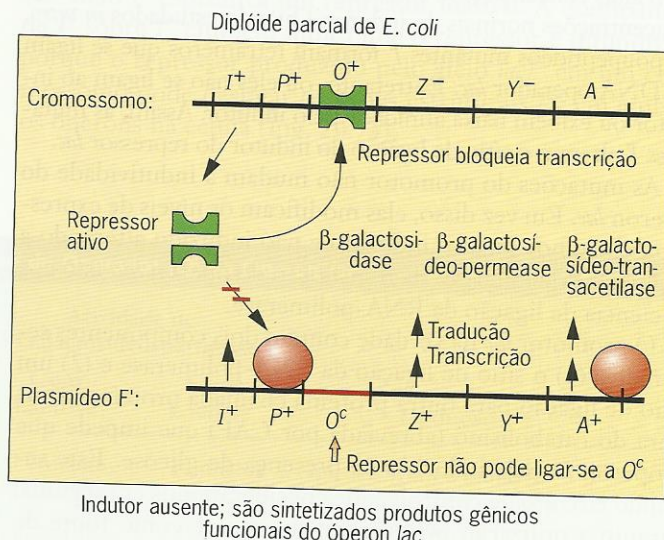
O Gene do Repressor *lac* (*I*) Atua em *Cis* e em *Trans*; o Operador *lac* Atua Apenas na Configuração *Cis*

| Genótipo                                       | Atividade de $\beta$ -Galactosidase <sup>a</sup> |              | Atividade de $\beta$ -Galactosídeo-permease <sup>a</sup> |              | Dedução                                               |
|------------------------------------------------|--------------------------------------------------|--------------|----------------------------------------------------------|--------------|-------------------------------------------------------|
|                                                | Com Lactose                                      | Sem Lactose  | Com Lactose                                              | Sem Lactose  |                                                       |
| $I^+ P^+ O^+ Z^+ Y^+$                          | 100 unidades                                     | 1 unidade    | 100 unidades                                             | 1 unidade    | Tipo selvagem é indutível                             |
| $I^+ P^+ O^+ Z^+ Y^+ / F' I^+ P^+ O^+ Z^- Y^-$ | 100 unidades                                     | 1 unidade    | 100 unidades                                             | 1 unidade    |                                                       |
| $I^- P^+ O^+ Z^+ Y^+ / F' I^+ P^+ O^+ Z^- Y^-$ | 100 unidades                                     | 1 unidade    | 100 unidades                                             | 1 unidade    | $I^+$ atua tanto em <i>cis</i> quanto em <i>trans</i> |
| $I^+ P^+ O^+ Z^+ Y^+ / F' I^+ P^+ O^c Z^- Y^-$ | 100 unidades                                     | 1 unidade    | 100 unidades                                             | 1 unidade    |                                                       |
| $I^+ P^+ O^c Z^+ Y^+ / F' I^+ P^+ O^c Z^- Y^-$ | 100 unidades                                     | 100 unidades | 100 unidades                                             | 100 unidades | $O^+$ atua apenas em <i>cis</i>                       |
|                                                |                                                  |              |                                                          |              | $O^c$ atua apenas em <i>cis</i>                       |

<sup>a</sup>Os níveis de atividade em bactérias tipo selvagem foram estabelecidos em 100 unidades tanto para  $\beta$ -galactosidase (o produto do gene *Z*) quanto para  $\beta$ -galactosídeo-permease (o produto do gene *Y*). O gene *A* e seu produto  $\beta$ -galactosídeo-transacetilase não são mostrados por motivo de abreviação.



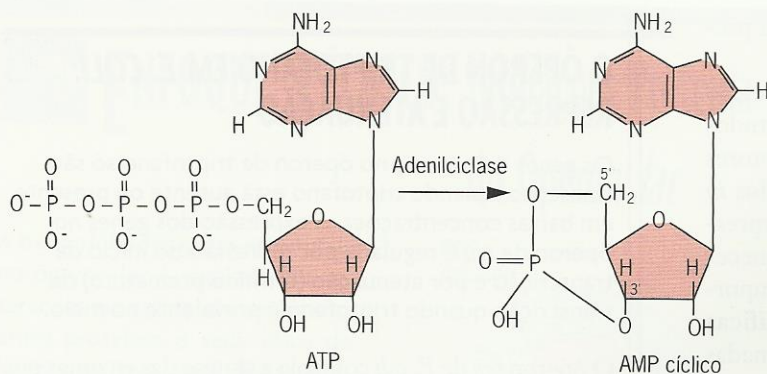
(a) Síntese indutível dos produtos gênicos do operon *lac* em uma bactéria  $F' I^+ P^+ O^c Z^- Y^- A^- / I^+ P^+ O^+ Z^+ Y^+ A^+$ .



(b) Síntese constitutiva dos produtos gênicos do operon *lac* em uma bactéria  $F' I^+ P^+ O^c Z^+ Y^+ A^+ / I^+ P^+ O^+ Z^- Y^- A^-$ .

**Fig. 20.8** ■ O operador atua apenas na configuração *cis*. A síntese de  $\beta$ -galactosidase funcional,  $\beta$ -galactosídeo-permease e  $\beta$ -galactosídeo-transacetilase é (a) indutível em um genótipo diplóide parcial  $F' I^+ P^+ O^c Z^- Y^- A^- / I^+ P^+ O^+ Z^+ Y^+ A^+$  e (b) constitutivo em um diplóide parcial de genótipo  $F' I^+ P^+ O^c Z^+ Y^+ A^+ / I^+ P^+ O^+ Z^- Y^- A^-$ . Estes resultados demonstram que o operador (*O*) é de ação *cis*, isto é, regula apenas os genes estruturais situados no mesmo cromossomo.





**Fig. 20.9** ■ A síntese de AMP cíclico (cAMP) catalisada por adenilciclase a partir de ATP.

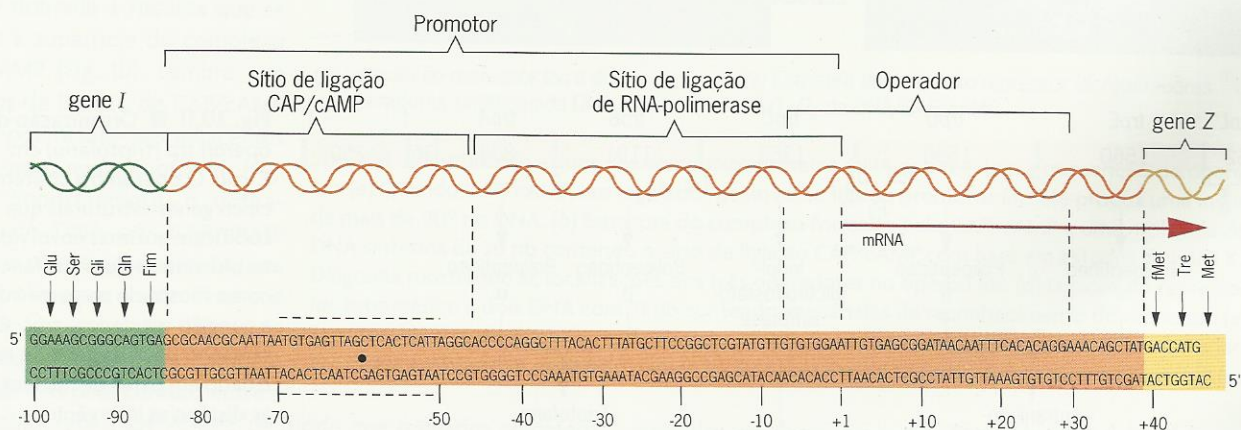
trolam enzimas envolvidas no catabolismo de carboidratos. Este fenômeno, chamado **repressão catabólica** (ou **efeito de glicose**), assegura que a glicose seja metabolizada quando presente, preferencialmente a outras fontes de energia menos eficientes.

A repressão catabólica do óperon *lac* e vários outros óperons é mediada por uma proteína reguladora chamada **CAP** (de **proteína ativadora do catabolismo**) e uma pequena molécula efetora chamada **AMP cíclico** (adenosina-3', 5'-monofosfato; abreviada por cAMP) (Fig. 20.9). Como a CAP se liga ao cAMP quando este mononucleotídeo está presente em concentrações suficientes, ela às vezes é chamada de proteína receptora de AMP cíclico.

O promotor *lac* contém dois sítios de ligação separados, um para RNA-polimerase e um para o complexo CAP/cAMP (Fig. 20.10). O complexo CAP/cAMP deve estar presente em seu sítio de ligação no promotor *lac* para que o óperon seja normalmente induzido. O complexo CAP/cAMP então exerce um controle positivo sobre a transcrição do óperon *lac*. Ele tem um efeito exatamente oposto ao da ligação do repressor a um operador. Embora o mecanismo exato pelo qual CAP/cAMP estimula a ligação da RNA-polimerase ao

promotor ainda seja incerto, seu controle positivo da transcrição do óperon *lac* está firmemente estabelecido pelos resultados de experimentos tanto *in vivo* quanto *in vitro*. CAP funciona como um dímero. Assim, como o repressor *lac*, ela é multimérica em seu estado funcional.

Apenas o complexo CAP/cAMP se liga ao promotor *lac*; na ausência de cAMP, CAP não se liga. Assim, cAMP atua como a molécula efetora, determinando o efeito de CAP na transcrição do óperon *lac*. A concentração intracelular de cAMP é sensível à presença ou à ausência de glicose. Altas concentrações de glicose causam acentuadas diminuições na concentração intracelular de cAMP. A glicose impede a ativação da adenilciclase, a enzima que catalisa a formação de cAMP a partir de ATP. Assim, a presença de glicose resulta em diminuição na concentração intracelular de cAMP. Na presença de baixa concentração de cAMP, CAP não pode ligar-se ao promotor do óperon *lac*. Por sua vez, a RNA polimerase não pode ligar-se eficientemente ao promotor *lac* na ausência de CAP/cAMP ligado. Logo, na presença de glicose, a transcrição do óperon *lac* nunca excede 2% da taxa induzida observada na ausência de glicose. Por mecanismos similares, CAP e cAMP impedem que sejam induzidos os



**Fig. 20.10** ■ Organização da região promotor-operador do óperon *lac*. O promotor consiste em dois componentes: (1) o sítio que liga o complexo CAP/cAMP e (2) o sítio de ligação da RNA-polimerase. Os segmentos adjacentes aos genes estruturais de *lacI* (repressor) e *lacZ* ( $\beta$ -galactosidase) também são mostrados. A linha horizontal no mRNA mostra a posição na qual começa a transcrição do óperon (a ponta 5' do mRNA *lac*). Os números na parte de baixo dão as distâncias em pares de nucleotídeos do sítio de início do transcrito (posição +1). O ponto (•) entre os dois filamentos de nucleotídeos indica o centro de simetria de um palíndromo imperfeito.



óperons de arabinose (*ara*) e galactose (*gal*) de *E. coli* na presença de glicose.

A sequência completa de pares de nucleotídeos da região reguladora do óperon *lac* é mostrada na Fig. 20.10. Estudos comparativos de sequências de nucleotídeos de promotores e operadores mutantes e tipo selvagem, além de estudos *in vitro* de ligação de CAP/cAMP, RNA-polimerase e repressor, bem como dados de cristalografia de raios X, forneceram informações detalhadas sobre a natureza destas importantes **interações proteínas-ácidos nucleicos específicas de sequências**. Algumas dessas interações são examinadas mais atentamente no Enfoque Técnico: Interações Proteína-DNA que Controlam a Transcrição do Óperon *lac*.

### PONTOS IMPORTANTES

- O óperon *lac* de *E. coli* é um sistema indutível negativo e catabólico repressível; os três genes estruturais no óperon *lac* são transcritos em níveis altos somente na presença de lactose e na ausência de glicose.
- Na ausência de lactose, o repressor *lac* liga-se ao operador *lac* e impede que a RNA-polimerase inicie a transcrição do óperon.
- A repressão catabólica impede que sejam induzidos óperons, tais como *lac*, que codificam enzimas envolvidas no catabolismo de carboidratos, na presença de glicose, a fonte de energia preferida.

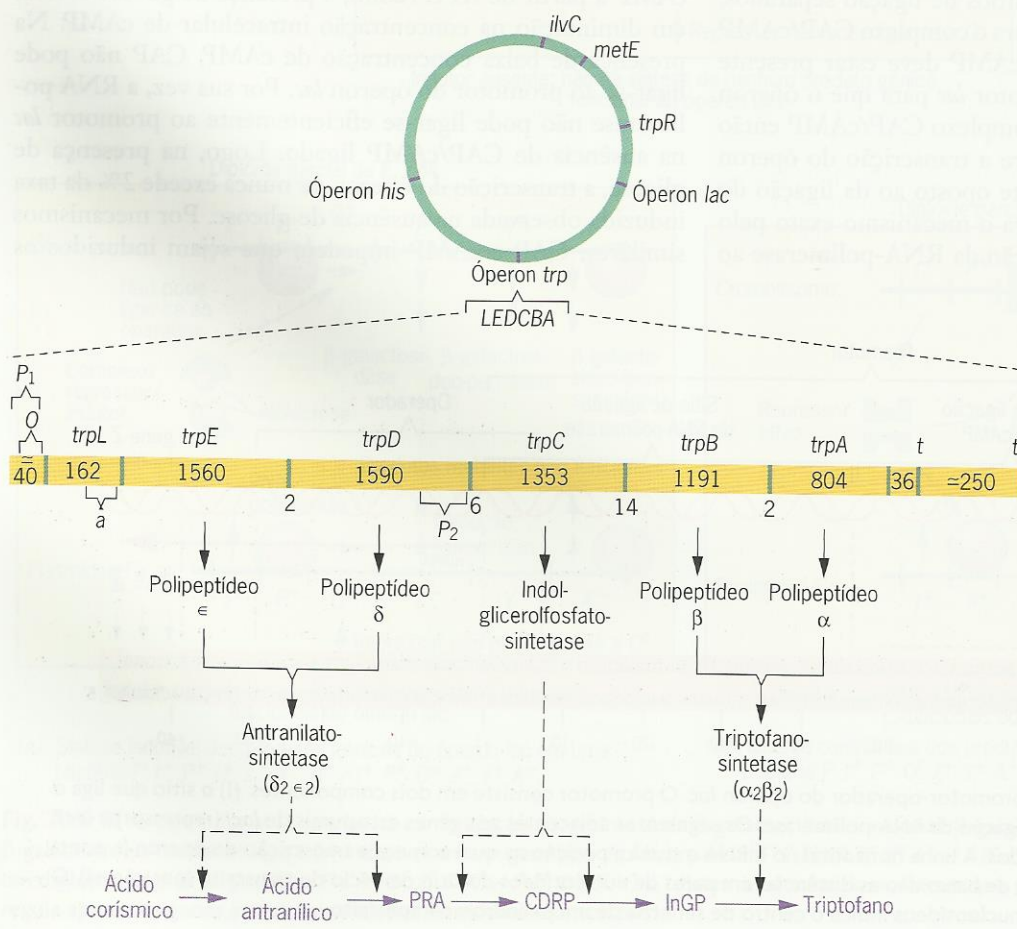
## O ÓPERON DE TRIPTOFANO EM *E. COLI*: REPRESSÃO E ATENUAÇÃO

Os genes estruturais no óperon de triptofano só são transcritos quando triptofano está ausente ou presente em baixas concentrações. A expressão dos genes no óperon de *trp* é regulada por repressão do início da transcrição e por atenuação (término prematuro) da transcrição quando triptofano é prevalente no meio.

O óperon *trp* de *E. coli* controla a síntese das enzimas que catalisam a biossíntese do aminoácido triptofano. As funções dos cinco genes estruturais e as sequências reguladoras adjacentes do óperon *trp* foram analisadas em detalhe por Charles Yanofsky e colaboradores. Os cinco genes estruturais codificam enzimas que convertem ácido corísmico em triptofano. A expressão do óperon *trp* é regulada em dois níveis: repressão, que controla o início da transcrição, e atenuação, que governa a frequência do término prematuro do transcrito. Discutiremos estes mecanismos reguladores nas duas seções seguintes.

### Repressão

O óperon *trp* de *E. coli* é um óperon repressível negativo. A organização do óperon *trp* e a via de biossíntese de triptofano são mostradas na Fig. 20.11. O gene *trp*, que codifica o



**Fig. 20.11** ■ Organização do óperon *trp* (triptofano) em *E. coli*. O óperon *trp* contém cinco genes estruturais que codificam enzimas envolvidas na biossíntese de triptofano, como mostrado na parte inferior, e a região reguladora *trpL*. O tamanho de cada gene ou região é dado em pares de nucleotídeos; as distâncias intergênicas são mostradas abaixo da sequência do gene. Chave: PRA, fosforribosil-antranilato; CDRP, carboxifenilamino-desoxirribulose fosfato; InGP, indol-glicerolfosfato.





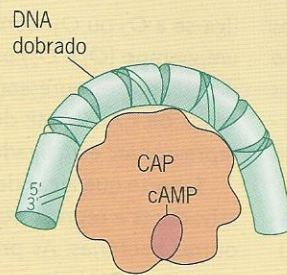
## ENFOQUE TÉCNICO: Interações Proteína-DNA que Controlam a Transcrição do Óperon *lac*

A transcrição dos genes estruturais no óperon *lac* é regulada por interações específicas de seqüências entre proteínas e seus sítios de ligação ao DNA. Uma dessas interações é a ligação da RNA-polimerase a seu sítio de ligação no promotor (discutida no Cap. 11). Uma segunda é a ligação de CAP/cAMP a seu sítio de ligação no promotor *lac* (discutida na seção anterior). Uma terceira é a ligação do repressor *lac* ao operador (na verdade, operadores).

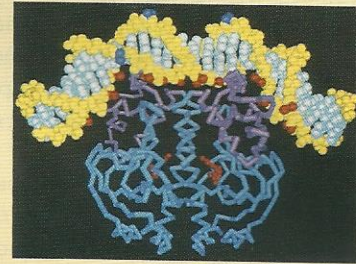
Vamos examinar primeiro a ligação de CAP/cAMP a seu sítio de ligação no promotor *lac*. CAP/cAMP controla a repressão catabólica; a ligação de CAP/cAMP ao promotor é necessária para indução eficiente do óperon *lac*. Como a ligação de CAP/cAMP estimula a transcrição dos genes estruturais *lac*? A RNA-polimerase não pode ligar-se eficientemente a seu sítio de ligação no promotor *lac* a menos que CAP/cAMP já esteja ligado. Quando CAP/cAMP se liga ao DNA, ele dobra o DNA (Fig. 1a). Estudos de raios X mostram que o DNA é dobrado à medida que se associa à superfície do complexo CAP/cAMP (Fig. 1b). Lembre que os sítios de ligação de CAP/cAMP e RNA-polimerase são adjacentes um ao outro no promotor *lac* (Fig. 20.10). Supostamente, o dobramento do DNA por CAP/cAMP promove um sítio mais aberto para a RNA-polimerase e assim favorece a ligação e transcrição dos genes estruturais. Entretanto, também há evidências para contato entre a

RNA-polimerase e CAP/cAMP, de modo que o quadro completo pode ser mais complexo do que apenas o dobramento do DNA.

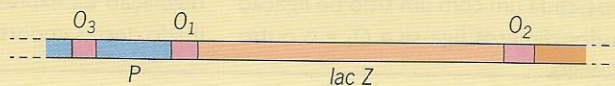
Em seguida, vamos examinar a ligação do repressor *lac* aos operadores, o que impede a RNA-polimerase de transcrever os genes estruturais no óperon. Originalmente, o óperon *lac* supostamente



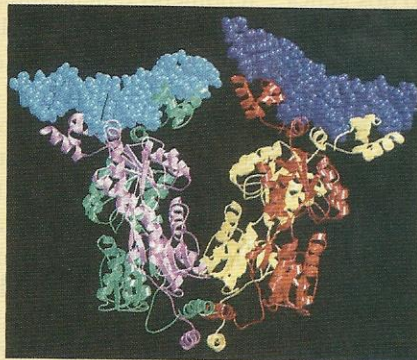
(a) DNA dobrado por CAP/cAMP.



(b) Estrutura do complexo CAP/cAMP/DNA.



(c) Localizações dos três operadores no óperon *lac*.



(d) Ligação do repressor *lac* a dois operadores sintéticos de DNA.



(e) Estrutura do complexo repressor *lac*/operadores  $O_1$ - $O_2$  de DNA/CAP/cAMP.

**Fig. 1 ■** Estruturas de alguns dos complexos proteína-DNA que regulam a transcrição do óperon *lac*. (a) Quando CAP/cAMP, um regulador positivo, se liga ao promotor *lac*, ele produz uma inclinação de mais de 90° no DNA. (b) Estrutura do complexo formado pelo CAP/cAMP e uma molécula de DNA sintética de 30 pb contendo o sítio de ligação CAP/cAMP com base em estudos de raios X. (c) Diagrama mostrando as localizações dos três operadores no óperon *lac*. (d) Ligação do repressor *lac* tetramérico a dois DNA com 21 pb contendo seqüências de reconhecimento do repressor. (e) Estrutura da alça de 93 pb formada quando o repressor tetramérico está ligado a operadores *lac*  $O_1$  e  $O_3$ . CAP/cAMP (azul) é mostrado dentro da alça associada a seu sítio de ligação no promotor *lac*.

continha um único operador. Posteriormente, foi demonstrado que ele contém três operadores, o operador primário,  $O_1$ , e dois operadores secundários,  $O_2$  e  $O_3$ .  $O_1$  é o operador original e está localizado entre o promotor e o gene Z (veja Fig. 20.10).  $O_2$  está situado posterior a  $O_1$  dentro do gene Z e  $O_3$  está situado ante-





## ENFOQUE TÉCNICO: Interações Proteína-DNA que Controlam a Transcrição do Óperon *lac* (continuação)

cedendo o promotor (Fig. 1c). Repressão máxima requer todos os três operadores. Entretanto, uma forte repressão ocorre enquanto  $O_1$  e  $O_2$  ou  $O_3$  estiverem presentes. Por que são necessários dois operadores para repressão eficiente? Para responder a esta pergunta, precisamos observar a ligação específica de sequência do repressor aos operadores.

A forma ativa do repressor *lac* é um tetrâmero contendo quatro cópias do produto do gene *lacI*. Estudos de raios X das estruturas formadas pelo repressor *lac* e sítios de ligação sintéticos com 21 pb mostraram que cada repressor tetramérico se liga a duas sequências operadoras simultaneamente (Fig. 1d). De fato, o tetrâmero consiste em dois dímeros, cada um com um sítio de ligação de sequência específica. Um dos dímeros liga-se a  $O_1$  e o outro

liga-se a  $O_2$  ou  $O_3$ . Com isso, o repressor dobra o DNA formando um grampo ( $O_1$  e  $O_2$ ) ou uma alça ( $O_1$  e  $O_3$ ). A estrutura proposta do complexo repressor- $O_1$ - $O_3$  é mostrada na Fig. 1e. Note a presença de CAP/cAMP dentro da alça de DNA formada quando o repressor *lac* está ligado tanto a  $O_1$  quanto a  $O_3$  (Fig. 1e).

Sabe-se que alças de DNA similares formam-se pela ligação de ativadores e repressores protéicos de outros óperons em *E. coli* e outras bactérias. As proteínas reguladoras têm a capacidade de se ligar ao DNA de modo específico de sequência, para alterar a estrutura do DNA e para estimular ou reprimir a transcrição de genes estruturais na vizinhança. Uma compreensão completa da regulação da expressão gênica irá requerer conhecimento detalhado destas importantes interações.

repressor *trp*, não está proximamente ligado ao óperon *trp*. A região do operador ( $O$ ) do óperon *trp* localiza-se dentro da região do promotor primário ( $P_1$ ). Há também um promotor fraco ( $P_2$ ) na ponta distal ao operador do gene *trpD*. O promotor  $P_2$  aumenta o nível basal de transcrição dos genes *trpC*, *trpB* e *trpA*. Duas sequências de término de transcrição ( $t$  e  $t'$ ) estão situadas posteriores a *trpA*. A região *trpL* especifica uma sequência líder com 162 nucleotídeos do mRNA.

A regulação da transcrição do óperon *trp* está diagramada na Fig. 20.4c. Na ausência de triptofano (o co-repressor), a RNA-polimerase liga-se à região do promotor e transcreve os genes estruturais do óperon. Na presença de triptofano, o complexo co-repressor/repressor liga-se à região do operador e impede que a RNA-polimerase inicie a transcrição dos genes no óperon.

A taxa de transcrição do óperon *trp* no estado desreprimido (ausência de triptofano) é 70 vezes a taxa que ocorre no estado reprimido (presença de triptofano). Em mutantes *trpR*, que não têm um repressor funcional, a taxa de síntese das enzimas biossintéticas de triptofano ainda está reduzida em cerca de 10 vezes pela adição de triptofano ao meio. Essa redução adicional na expressão do óperon *trp* é causada por atenuação, que é discutida a seguir.

### Atenuação

Deleções que removem parte da região *trpL* (Fig. 20.11) resultam em taxas aumentadas de expressão do óperon *trp*. Entretanto, essas deleções não têm efeito na repressibilidade do óperon *trp*. Isto é, repressão e desrepressão ocorrem como em linhagens *trpL*<sup>+</sup>. Tais resultados indicam que a síntese das enzimas de biossíntese de triptofano é regulada em um segundo nível por um mecanismo que é independente de

repressão/desrepressão e requer sequências de nucleotídeos presentes na região *trpL* do óperon *trp*.

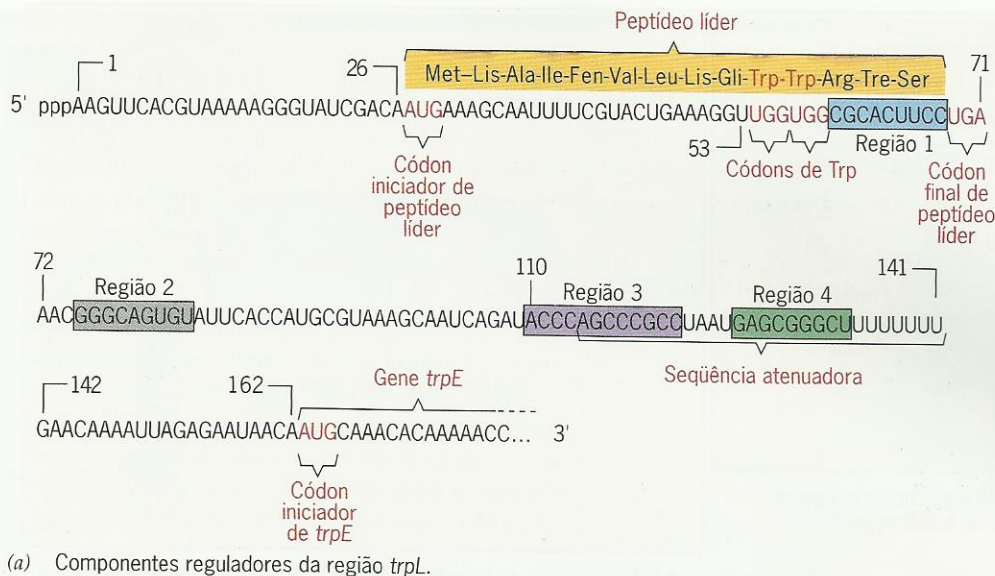
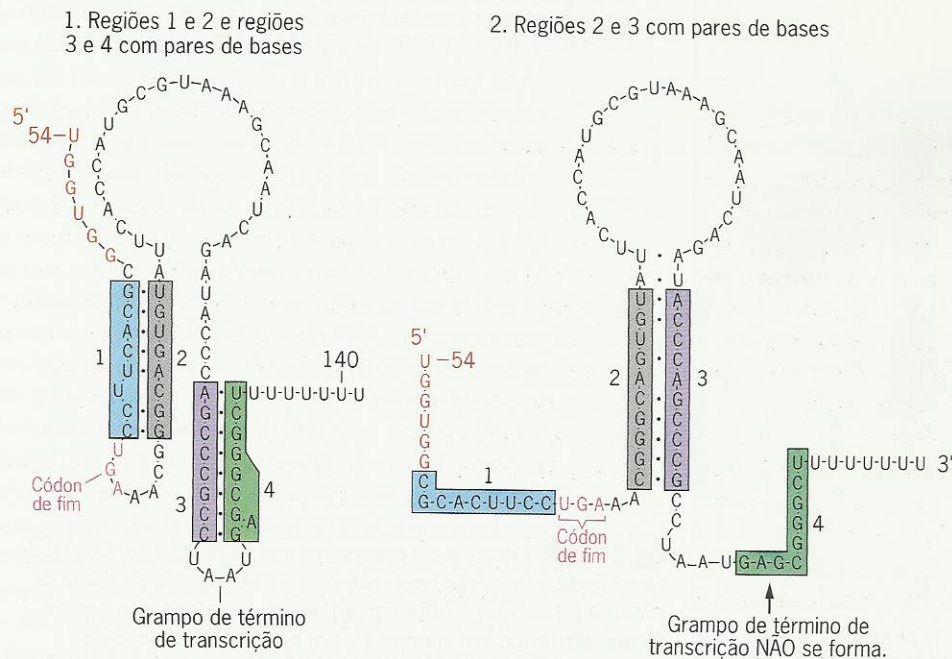
Este segundo nível de regulação do óperon *trp* é chamado de **atenuação**, e a sequência dentro de *trpL* que controla este fenômeno é denominada **atenuador** (Fig. 20.12a). Ocorre atenuação por controle do término da transcrição em um sítio perto da ponta da sequência líder do mRNA. Esse término "premature" da transcrição do óperon *trp* ocorre apenas na presença do tRNA<sup>Trp</sup> carregado com triptofano. Quando ocorre este término prematuro ou atenuação, é produzido um transcrito *trp* (140 nucleotídeos).

A região atenuadora tem uma sequência de pares de nucleotídeos essencialmente idêntica aos **sinais de término de transcrição** encontrados nas pontas da maioria dos óperons bacterianos. Esses sinais de término contêm um palíndromo rico em G:C seguido de vários pares de bases A:T. A transcrição destes sinais de término produz um RNA nascente com o potencial de formar uma estrutura em grampo com pontes de hidrogênio seguida de várias uracilas. Quando um transcrito nascente forma esta estrutura em grampo, ela causa uma mudança conformacional na RNA-polimerase associada, resultando em término da transcrição dentro da região seguinte, com fracas pontes de hidrogênio (A:U)<sub>n</sub> de pareamentos de base DNA-RNA.

A sequência de nucleotídeos do atenuador explica portanto sua capacidade de terminar prematuramente a transcrição do óperon *trp*. Mas como isto pode ser regulado pela presença ou pela ausência de triptofano?

Primeiro, lembre que transcrição e tradução são acopladas em procariontes, isto é, os ribossomos começam a traduzir mRNA enquanto eles ainda estão sendo formados. Assim, eventos que ocorrem durante a tradução também podem afetar a transcrição.



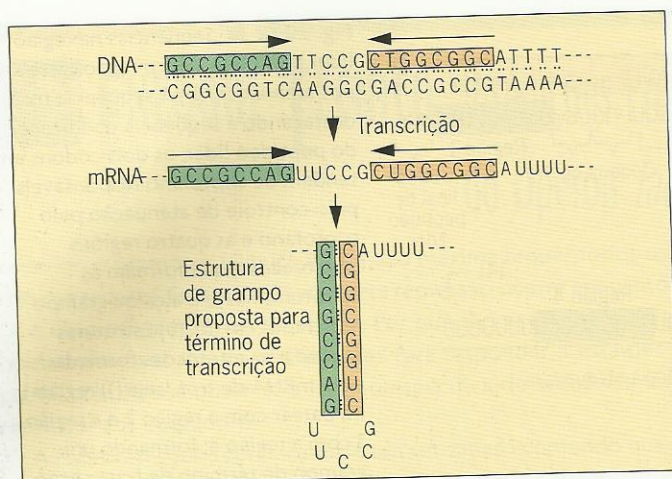
(a) Componentes reguladores da região *trpL*.(b) Estruturas secundárias alternadas formadas pelo transcrito *trpL*.

**Fig. 20.12** ■ Sequências na região líder do mRNA de *trp* responsáveis por atenuação. (a) A sequência *trpL*, destacando a sequência codificante do peptídeo líder, os dois códons em tandem de triptofano responsáveis pelo controle de atenuação pelo triptofano e as quatro regiões (sombreadas) que formam as estruturas haste e alça ou grampo mostradas em (b). (b) Estruturas secundárias alternadas formadas pelo mRNA de *trpL*, seja (1) região 1 irá parear com a região 2 e a região 3 com a região 4, formando um grampo de término de transcrição, ou (2) a região 2 fará par de bases com a região 3, impedindo que a região 3 faça par com a região 4. A concentração de triptofano na célula determina qual destas estruturas irá formar-se durante a transcrição do operon *trp*.

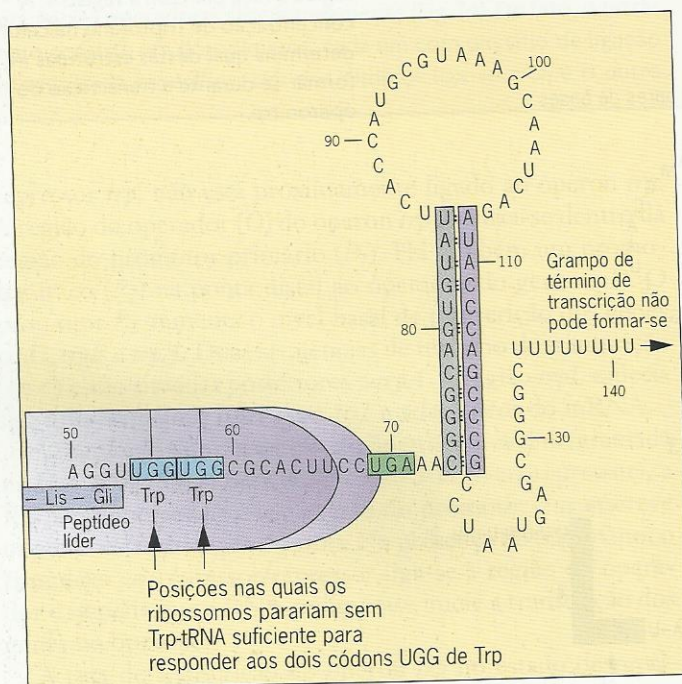
Segundo, note que a sequência líder com 162 nucleotídeos do mRNA do operon *trp* contém sequências que podem fazer pares de bases para formar estruturas alternadas haste e alça ou grampo (Fig. 20.12b). As quatro regiões líderes que podem fazer pares de bases para formar estas estruturas são: (1) nucleotídeos 60–68, (2) nucleotídeos 75–83, (3) nucleotídeos 110–121 e (4) nucleotídeos 126–134. Os tamanhos destas regiões envolvidas no pareamento de bases variam dependendo de qual região está pareada. As sequências de nucleotídeos destas quatro regiões são tais que a região 1 pode fazer par de bases com a região 2, a região 2 pode parear com a região 3 e a região 3 pode parear com a região 4. A região 2 pode parear

ou com a região 1 ou com a 3, mas, obviamente, só pode parear com uma destas regiões de cada vez. Assim, existem duas possíveis estruturas secundárias para a sequência líder *trp*: (1) região 1 pareada com região 2 e região 3 pareada com região 4 ou (2) região 2 pareada com região 3, deixando as regiões 1 e 4 não pareadas. O pareamento das regiões 3 e 4 produz o já mencionado grampo de término de transcrição. Se a região 3 faz pares de bases com a região 2, ela não pode parear com a região 4, e o grampo de término de transcrição não pode ser formado. Como você provavelmente já imaginou, a presença ou a ausência de triptofano determina quais destas alternativas estruturais irá formar-se.





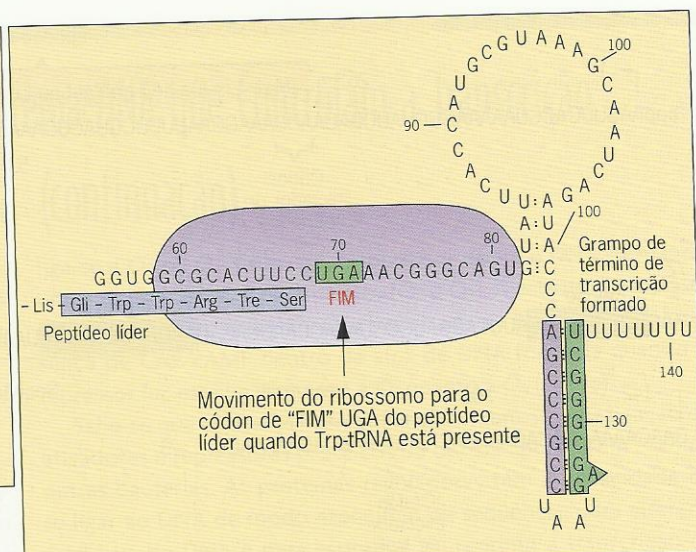
(a) Estrutura da sequência 5' do término da transcrição do operon *trp* e formação do grampo de término de transcrição.



(b) Com baixos níveis de triptofano, a tradução da sequência líder pára em um dos códons de Trp. Essa parada permite que as regiões líderes 2 e 3 formem par, o que impede a região 3 de parear com a região 4 para formar o grampo de término de transcrição. Assim, a transcrição continua por todo o operon *trp*.

Terceiro, note que a sequência líder contém um códon AUG de início de tradução, seguido de 13 códons para aminoácidos e seguido por sua vez de um códon de término de tradução UGA (Fig. 20.12a). Além disso, a sequência líder *trp* contém um eficiente sítio de ligação de ribossomo situado na posição apropriada para o início da tradução do códon AUG líder de iniciação. Todas as evidências disponíveis indicam que um "peptídeo líder" com 14 aminoácidos é sintetizado, conforme diagramado na Fig. 20.12a.

O grampo de término de transcrição do operon normal *trp* é mostrado na Fig. 20.13a, e o mecanismo proposto de



(c) Na presença de triptofano suficiente, a tradução continua além dos códons de Trp até o códon de término e perturba o pareamento de bases entre as regiões líder 2 e 3. Este processo deixa a região 3 livre para parear com a região 4 e formar o grampo de término de transcrição, que pára a transcrição na sequência atenuadora.

**Fig. 20.13** ■ Controle do operon *trp* por atenuação. (a) O sinal de término de transcrição em *E. coli* contém uma região de diade de simetria (setas) que resulta em sequências de mRNA que podem formar estruturas em grampo. (b) Em baixa concentração de triptofano, a transcrição continua através da sequência atenuadora por todo o operon *trp*. (c) Na presença de triptofano suficiente, a transcrição frequentemente termina na sequência atenuadora.

atenuação da transcrição do operon *trp* está diagramado na Fig. 20.13b e c. O peptídeo líder contém dois triptofanos contíguos. Os dois códons Trp estão posicionados de tal modo que, em baixas concentrações de triptofano (e assim baixas concentrações de Trp-tRNA<sup>Trp</sup>), o ribossomo iria parar antes de encontrar a estrutura de bases pareadas formada pelas regiões líder 2 e 3 (Fig. 20.13b). Como o pareamento das regiões 2 e 3 impede a formação de grampos de término de transcrição por pareamentos de bases das regiões 3 e 4, as transcrições continuarão após o atenuador no gene *trpE* na ausência de triptofano.





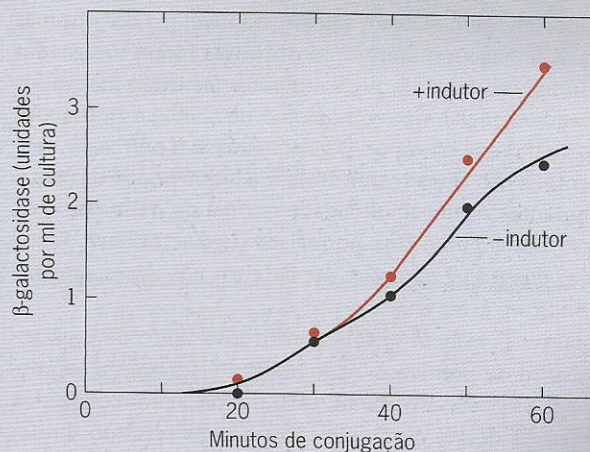
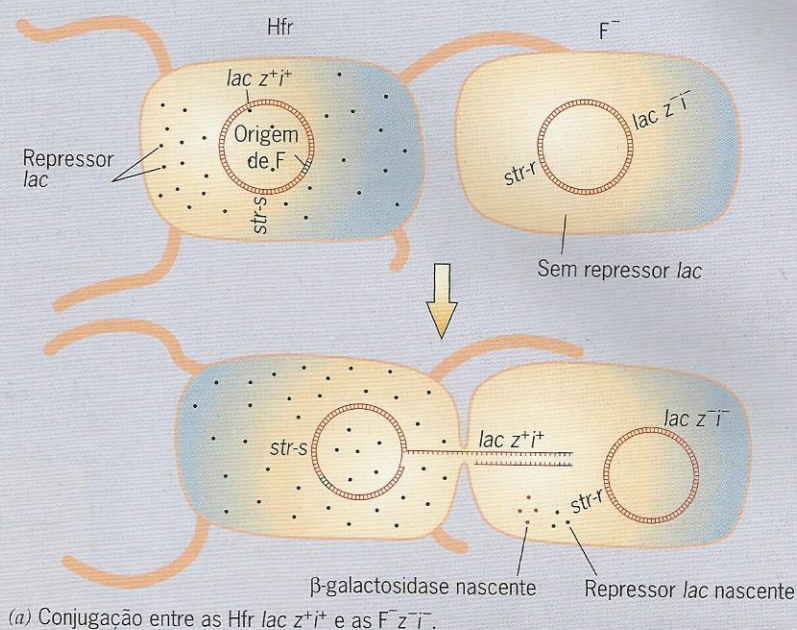
## UM MARCO NA GENÉTICA: Jacob, Monod e o Modelo do Óperon

Aqui estou hoje, compartilhando com André Lwoff e Jacques Monod esta grande honra que nos foi concedida, sem dúvida porque, quando comecei uma pesquisa em 1950, tive a sorte de estar no lugar certo na época certa. No lugar certo, porque lá, no Instituto Pasteur, estava surgindo uma nova disciplina em uma atmosfera de entusiasmo, críticas lúcidas, não conformismo e amigável. No momento certo, porque a biologia borbulhava de atividade, mudando seus modos de pensar, descobrindo em microrganismos um material novo e simples e seguindo junto à física e à química. Um raro momento, no qual o ignorado podia ser uma virtude.<sup>1</sup>

Este pronunciamento introdutório feito por François Jacob ao receber o Prêmio Nobel dá uma visão informativa sobre uma era especial na história da genética molecular. Por todos os relatos, o Instituto Pasteur desta época era um ambiente especial que favorecia uma livre troca de idéias, juntamente com uma avaliação crítica de seus méritos e deficiências. Foi nas instalações do Instituto Pasteur que Jacob e Monod desenvolveram o modelo do óperon para explicar os resultados de seus estudos sobre o óperon *lac* de *E. coli*.<sup>2</sup>

Quando Jacob mencionou “a época certa”, ele se referia em parte à descoberta do RNA mensageiro que leva informação genética dos genes para os ribossomos. Em 1961, Sydney Brenner, Jacob e Matthew Meselson demonstraram a síntese de proteínas do fago T4 em ribossomos que estavam presentes em *E. coli* antes da infecção.<sup>3</sup> Seus resultados proporcionaram evidências para uma molécula de RNA de vida curta que medeia a síntese de proteínas. O modelo do óperon não teria tido sentido na ausência desse RNA intermediário instável.

Quando Jacob e Monod começaram seu estudo sobre utilização de lactose em *E. coli*, eles observaram que a  $\beta$ -galactosidase era sintetizada apenas na presença do substrato lactose ou galactosídeos proximalmente relacionados. Seu colega André Lwoff os informou que provavelmente estavam observando uma “adaptação enzimática” e lhes indicou publicações sobre o assunto. De fato, este fenômeno foi descoberto por Émile Duclaux em 1899<sup>4</sup> e amplamente estudado durante os 60 anos se-



**Fig. 1** ■ Síntese de  $\beta$ -galactosidase nos diplóides parciais transitórios produzidos em cruzamentos entre bactérias Hfr *lac z<sup>+</sup> i<sup>+</sup>* e *F<sup>-</sup> lac z<sup>-</sup> i<sup>-</sup>*. (a) Diagrama do cruzamento. (b) Síntese de  $\beta$ -galactosidase nos diplóides parciais na presença e ausência de indutor.

guintes. Embora as mudanças nos níveis enzimáticos fossem bem caracterizadas, o mecanismo subjacente era desconhecido.

Jacob e Monod isolaram linhagens mutantes de *E. coli* com defeitos em dois genes estruturais importantes no locus *lac*. Essas mutações foram designadas como *z<sup>-</sup>* (perda de  $\beta$ -galactosidase) e *y<sup>-</sup>* (perda de  $\beta$ -galactosídeo-permease).<sup>5</sup> Eles também isola-

<sup>5</sup> Como o modelo original de Jacob e Monod é reproduzido na Fig. 2, usaremos aqui seus símbolos genéticos em letra minúscula, em vez dos símbolos com letras maiúsculas atualmente usados (veja texto).

<sup>1</sup>Jacob, F. 1965. Genetics of the bacterial cell. Em *Noble Lectures in Molecular Biology*, 1933-1975, pp. 219-244, Elsevier, New York.

<sup>2</sup>Jacob, F., e J. Monod. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3:318-356.

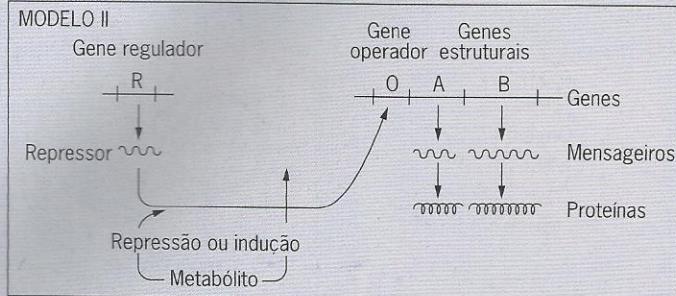
<sup>3</sup>Brenner, S., F. Jacob e M. Meselson. 1961. An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature* 190: 576-580.

<sup>4</sup>Duclaux, E. 1899. *Traité de Microbiologie*. Masson et Cie, Paris.





## UM MARCO NA GENÉTICA (continuação)



**Fig. 2** ■ Diagrama do modelo do óperon como aparece na publicação de Jacob e Monod de 1961. No modelo II (não mostrado), o repressor interage com o mRNA e impede sua tradução. Em sua publicação, Jacob e Monod referem-se ao operador como um gene, o que não fazemos hoje em dia.

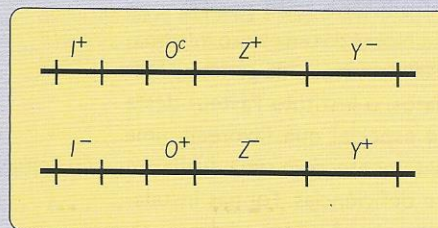
ram mutantes nos quais a síntese de  $\beta$ -galactosidase e permease ocorre continuamente, independente de lactose estar presente. Eles chamaram este fenômeno de *síntese constitutiva* e referiram as linhagens que a exibiam como *mutantes constitutivos*. As linhagens tipo selvagem exibiam síntese indutível ( $i^+$ ) das enzimas *lac*. Assim, as linhagens mutantes constitutivas foram designadas como  $i^-$ . Todos os vários tipos de mutantes *lac* estão mapeados em uma pequena região do cromossomo de *E. coli*. As mutações  $z^-$  e  $y^-$  resultaram na perda de  $\beta$ -galactosidase e permease, respectivamente, e não tinham efeito sobre a indutibilidade das enzimas. Os mutantes  $z^-$  tinham atividade normal de permease e os mutantes  $y^-$  tinham atividade normal de  $\beta$ -galactosidase. Tanto mutações  $z^-$  quanto  $y^-$  pareciam ser defeitos nos genes estruturais que codificam estas enzimas.

Os mutantes  $i^-$  exibem síntese constitutiva tanto de  $\beta$ -galactosidase quanto de permease. As mutações  $i^-$  alteram a expressão de ambos os genes estruturais. Uma das primeiras conquistas na compreensão destes mutantes constitutivos resultou de um trabalho colaborativo com Arthur Pardee, um membro de pós-doutorado naquela época. Ele construiu linhagens Hfr e F<sup>-</sup> que tinham várias combinações de marcadores *lac* e então fez experimentos de cruzamento com estas linhagens. Um cruzamento era entre uma linhagem  $z^+ i^+ \text{Hfr}$  com uma linhagem  $z^- i^- \text{F}^-$ . As Hfr não podiam produzir  $\beta$ -galactosidase na ausência do indutor lactose porque algum tipo de repressor estava presente, e as F<sup>-</sup> não podiam sintetizar esta enzima porque tinham um defeito no gene estrutural. Entretanto, durante a reprodução, os genes  $z^+$  das Hfr foram expressos imediatamente nos diplóides parciais transitórios formados por transferência do gene  $z^+$  para as F<sup>-</sup> (Fig. 1).<sup>6</sup> Como outros experimentos mostraram que  $i^+$  é dominante em relação a

$i^-$ , este resultado demonstrou que o fenótipo  $i^+$  versus  $i^-$  depende do estado do citoplasma.

Embora os diplóides parciais transitórios forneçam resultados informativos, sua instabilidade impede uma melhor análise das várias mutações *lac*. O problema da instabilidade foi logo eliminado pela descoberta dos fatores F que levam segmentos do cromossomo de *E. coli* (fatores F'; veja Cap. 8).<sup>7</sup> Jacob e Monod usaram fatores F' que levam genes *lac* para produzir diplóides parciais estáveis que continham mutações *lac* em muitas combinações diferentes. Um exemplo é discutido adiante.

Além de mutações constitutivas  $i^-$  mencionadas anteriormente, Jacob e Monod identificaram mutações constitutivas que foram mapeadas mais próximas ao gene *z* que as mutações  $i^-$ . Ao contrário dos mutantes  $i^-$ , estes mutantes constitutivos,



chamados mutantes  $o^c$ , de operador constitutivo, eram dominantes em relação a seu alelo tipo selvagem (indutível); isto é, os genes estruturais

*lac* foram expressos constitutivamente em diplóides parciais  $o^c z^+ y^+ / o^+ z^+ y^+$ . Na verdade, os alelos  $o^c$  são *cis*-dominantes. Eles só afetam a expressão de genes situados em *cis*, no mesmo cromossomo, a si próprios. Assim, um diplóide parcial do tipo produz  $\beta$ -galactosidase (produto do gene *z*) constitutivamente e  $\beta$ -galactosídeo-permease (produto do gene *y*) indutivelmente. Este resultado indicou a Jacob e Monod que o gene regulador  $i^+$  atua tanto em *cis* quanto em *trans* e codifica um produto difundível. Muito embora o gene  $i^+$  esteja no cromossomo de cima no diagrama, seu produto, um repressor, difunde-se pela célula, liga-se a  $o^+$  no cromossomo de baixo e impede a transcrição dos genes neste cromossomo quando o indutor lactose está ausente. Em contraste, o operador **não** faz um produto difundível. Ele é o sítio de ligação do repressor. Assim,  $o^c$  e  $o^+$  só afetam a transcrição de genes estruturais situados em *cis* a si mesmos. No diplóide parcial citado,  $o^c$  controla a expressão de *lac z*<sup>+</sup> e  $o^+$  controla a expressão de *lac y*<sup>+</sup>. Assim, o modelo do óperon de Jacob e Monod distingue entre os elementos reguladores de ação *cis* e *trans*.

A clareza e precisão do modelo de óperon de Jacob e Monod foi incrível. Seu modelo original é reproduzido na Fig. 2. Naquela época, eles não sabiam se a repressão ocorria em nível da transcrição (Modelo I, mostrado na Fig. 2) ou em nível da tradução (Modelo II, não mostrado). Eles também concluíram incorreta-

<sup>6</sup>Pardee, A. B., F. Jacob e J. Monod. 1959. The genetic control and cytoplasmic expression of "inducibility" in the synthesis of  $\beta$ -galactosidase by *E. coli*. *J. Mol. Biol.* 1:165-178.

<sup>7</sup>Jacob, F., e E. A. Adelberg. 1959. Transfert de caractères génétiques par incorporation au facteur sexuel d'*Escherichia coli*. *Comptes Rendus des Séances de L'Académie des Sciences* 249:189-191. Tradução para o inglês publicada em *Papers on Bacterial Genetics* (E. A. Adelberg, ed.), 1960. Little, Brown, Boston.





## UM MARCO NA GENÉTICA (continuação)

mente que o repressor era RNA, e não proteína. Esta conclusão foi baseada em experimentos nos quais 5-metiltryptofano (5-mT) foi usado para inibir a síntese de proteínas. Síntese do repressor ocorreu na presença de 5-mT. Assim, Jacob e Monod concluíram que o repressor não podia ser uma proteína. Pesquisas subsequentes mostraram que 5-mT bloqueia a síntese de proteínas apenas parcialmente e que o repressor *lac* é sintetizado em sua presença.

Em 1965, apenas alguns anos após terem proposto o modelo do operon, Jacob e Monod compartilharam o Prêmio Nobel de Fisiologia ou Medicina com André Lwoff.<sup>8</sup> Obviamente, o mundo científico logo reconheceu o significado do trabalho de Jacob e Monod. De fato, por muitos anos, o modelo do operon foi o paradigma para a regulação da expressão gênica. Apenas quando pesquisadores começaram a focar os eucariontes é que outros mecanismos reguladores passaram a ser a fronteira.

<sup>8</sup>Lwoff recebeu parte do Prêmio Nobel por seu trabalho sobre a correlação entre bactérias lisogênicas e bacteriófagos temperados (veja Cap. 8).

Na presença de triptofano suficiente, o ribossomo pode traduzir além dos códons Trp até o códon de término do peptídeo líder. No processo, ele irá perturbar o pareamento de bases entre as regiões líderes 2 e 3. Essa perturbação deixa a região 3 livre para parear com a região 4, formando o grampo de término de transcrição (Fig. 20.13c). Assim, na presença de triptofano suficiente, a transcrição frequentemente (cerca de 90% das vezes) termina no atenuador, reduzindo a quantidade de mRNA para os genes estruturais *trp*.

A transcrição do operon *trp* pode ser regulada em cerca de quase 700 vezes pelos efeitos combinados de repressão (até 70 vezes) e atenuação (até 10 vezes).

Regulação da transcrição por atenuação não é única do operon *trp*. Sabe-se que cinco outros operons (*trc*, *ilv*, *leu*, *fen* e *his*) são regulados por atenuação. O operon *his*, que por muitos anos se acreditou ser repressível, hoje é considerado regulado totalmente por atenuação. Embora pequenos detalhes variem de operon para operon, as características principais de atenuação são as mesmas para todos os seis operons.

### PONTOS IMPORTANTES

- O operon *trp* de *E. coli* é um sistema repressível negativo; a transcrição dos cinco genes estruturais no operon *trp* é reprimida na presença de concentrações significativas de triptofano.
- Operons tais como *trp* que codificam enzimas envolvidas nas vias de biossíntese de aminoácidos em geral são controlados por um segundo mecanismo regulador chamado atenuação.

### QUESTÕES PARA DISCUSSÕES

1. Quando Jacob e Monod desenvolveram o modelo do operon no final dos anos 1950 e início dos anos 1960, o Instituto Pasteur onde eles trabalharam era um ambiente que estimulava a livre troca de informações e idéias. Como a livre troca de idéias e informações resultou no desenvolvimento do modelo do operon? Essa livre troca é importante para o progresso da ciência?
2. O modelo do operon teve um grande impacto sobre como os geneticistas pensam sobre a regulação da expressão gênica. Por vários anos, geneticistas pensaram que os operons de histidina de *E. coli* e *Salmonella typhimurium* eram regulados por repressão, como o operon *lac*. Entretanto, estudos nos anos 1970 mostraram que esses operons são regulados por atenuação, como o operon de triptofano discutido neste capítulo. Claramente, cientistas, como qualquer outro indivíduo, são influenciados por paradigmas familiares. Se você inicia uma carreira em ciência, como você compensa esta influência? O que você pode fazer para garantir que sua mente esteja aberta a todas as explicações possíveis de um determinado resultado?

- Atenuação ocorre pelo término prematuro da transcrição em um sítio na sequência líder do mRNA (a sequência a 5' da região codificante) quando triptofano é prevalente no meio em que a bactéria está sendo cultivada.

## BACTERÍÓFAGO LAMBDA: LISOGENIA OU LISE

Os bacteriófagos temperados, tais como o fago  $\lambda$ , podem (1) entrar na via lítica, durante a qual produzem grandes números de fagos e matam a célula hospedeira, ou (2) entrar na via lisogênica, durante a qual inserem covalentemente seu cromossomo, em um estado "reprimido" ou dormente, no cromossomo do hospedeiro.

Quando um bacteriófago temperado tal como lambda ( $\lambda$ ) infecta uma bactéria, ele pode seguir uma dentre duas vias de desenvolvimento (veja Fig. 8.5). Ele pode ou (1) entrar no ciclo lítico, durante o qual ele se reproduz e lisa a célula hospedeira, como um fago virulento (Cap. 8, veja Fig. 8.3), ou (2) entrar na via lisogênica (Cap. 8, veja Fig. 8.6), durante a qual seu cromossomo é inserido no cromossomo do hospedeiro e se replica como qualquer outro segmento deste cromossomo. Quando integrado ao cromossomo da célula hospedeira, o cromossomo do fago é chamado *pro-*



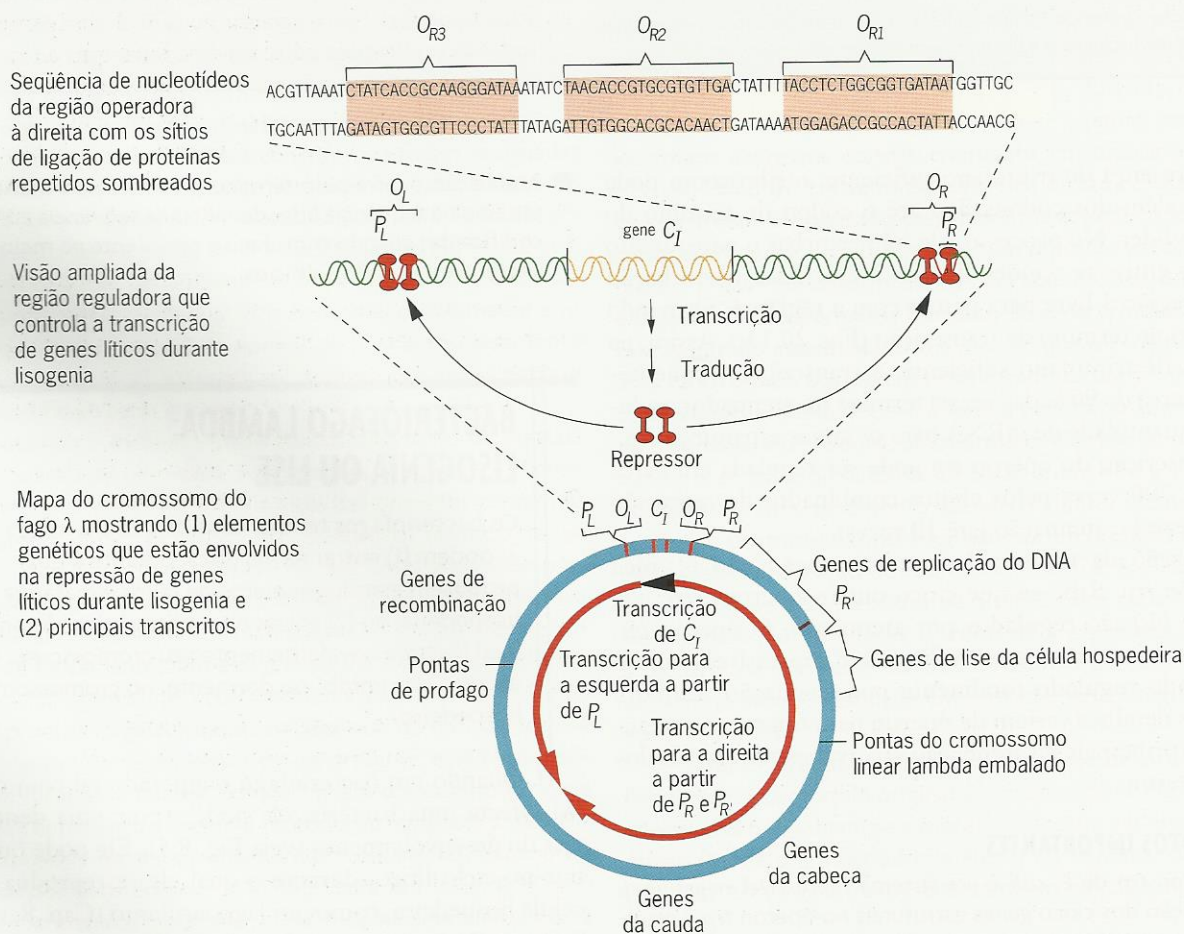
*fago*. Em uma bactéria lisogênica, os genes do profago que codificam produtos envolvidos na via lítica não devem ser expressos. Os genes do profago que especificam enzimas envolvidas na replicação do DNA do fago, proteínas estruturais necessárias para a morfogênese do fago e a lisozima que catalisa a lise da célula devem ser mantidos desligados para manter um estado lisogênico estável.

Os genes da via lítica do profago são mantidos desligados nas células lisogênicas por um simples circuito repressor-operador-promotor, muito similar aos circuitos reguladores negativos dos óperons de bactérias. Entretanto, o aspecto realmente interessante do fago lambda é o mecanismo que determina se  $\lambda$  entra na via lítica ou na via lisogênica. Essa determinação envolve uma mudança genética elegante que direciona o cromossomo infectante do fago para uma dentre duas vias: (1) a cascata reguladora lítica ou (2) o circuito de repressão mantido autogenamente de lisogenia. O intercâmbio destas duas redes reguladoras é uma das mais fascinantes histórias no campo da genética molecular. No final, se uma proteína, o repressor  $\lambda$ , ocupa sítios reguladores importantes, ocorre lisogenia. Se outra proteína  $\lambda$ , chamada Cro, ocupa

esses sítios, ocorre desenvolvimento lítico. Entretanto, antes de discutir a complexa mudança lise *versus* lisogenia, precisamos conhecer alguns genes reguladores de lambda.

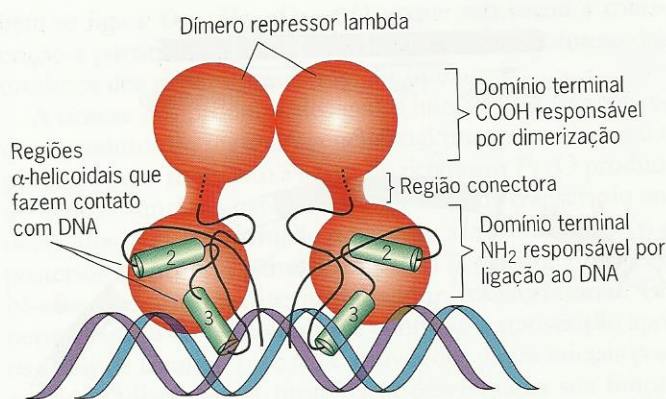
## Repressão de Genes da Via Lítica Lambda Durante Lisogenia

Os genes do fago lambda que codificam funções envolvidas no desenvolvimento lítico são mantidos desligados por um mecanismo de controle negativo análogo aos que controlam os óperons *lac* e *trp* de *E. coli*. O gene  $C_I$  do fago  $\lambda$  codifica um repressor, uma proteína bem caracterizada com um peso molecular de 27.000 daltons. O repressor  $C_I$ , como um dímero, liga-se a duas regiões operadoras que controlam a transcrição de genes  $\lambda$  envolvidos em crescimento lítico (Fig. 20.14). Estas duas regiões operadoras, chamadas de  $O_L$  (para transcrição no sentido da esquerda) e  $O_R$  (para transcrição no sentido da direita), ficam superpostas a seqüências promotoras às quais a RNA-polimerase se liga para iniciar a transcrição de genes que controlam o desenvolvimento



**Fig. 20.14** ■ Repressão dos genes líticos lambda em uma *E. coli* lisogênica. A transcrição dos genes líticos é reprimida pela ligação do repressor lambda a duas seqüências operadoras ( $O_L$  e  $O_R$ ), que regulam a transcrição para a esquerda e para a direita do cromossomo lambda (em baixo). O repressor lambda é codificado pelo gene  $C_I$  (centro) e reprime a síntese dos principais transcritos ligando-se a triplos sítios de ligação da proteína em  $O_L$  e  $O_R$ . As setas mostram os tamanhos relativos e direções de síntese dos principais transcritos lambda.





**Fig. 20.15** ■ O dímero de repressor lambda contata seus sítios de ligação nas regiões do operador lambda ( $O_L$  e  $O_R$ ) primariamente por meio de duas regiões  $\alpha$ -helicoidais (cilindros verdes) no domínio aminoterminal de ligação ao DNA. Embora a conformação exata do domínio de terminal carboxila seja desconhecida, a estrutura tridimensional (linha preta e cilindros verdes) do domínio aminoterminal já foi determinada. Este domínio de ligação do DNA contém cinco regiões  $\alpha$ -hélice, duas das quais (marcadas como 2 e 3 no diagrama) são primariamente responsáveis pela ligação específica do repressor às regiões do operador lambda. (Baseada em Benjamin Lewin, *Genes V*, Cell Press, 1994. Usada com permissão.)

lítico. Com o repressor ligado a estes dois operadores, a RNA-polimerase não pode ligar-se aos dois promotores e não pode iniciar a transcrição. Deste modo, os líticos de lambda continuam reprimidos, permitindo que o profago dormente seja transmitido de células hospedeiras parentais para a geração de células prole geração após geração.

Cada operador lambda contém três sítios de ligação de repressor, cada um com 17 pares de nucleotídeos de tamanho, com seqüências de nucleotídeos similares mas não idênticas. Cada sítio de ligação de repressor tem uma simetria parcial dupla ao redor do par de base central. Essa simetria tem um papel importante na interação de dímeros de repressores com tais sítios de ligação de operador (Fig. 20.15).

O repressor lambda foi estudado em grande detalhe. O monômero do repressor contém três regiões: (1) um **domínio de ligação do DNA** no terminal amino, (2) um **domínio de dimerização** no terminal carboxila e (3) uma **região conectora** central. A estrutura tridimensional é conhecida apenas para o domínio de ligação do DNA do repressor. Esse domínio, com 92 aminoácidos de tamanho, contém cinco regiões que formam estruturas de  $\alpha$ -hélice; as estruturas de  $\alpha$ -hélice são numeradas de 1 a 5, começando no terminal  $NH_3$ . Duas das regiões da  $\alpha$ -hélice, números 2 e 3, são amplamente responsáveis pela especificidade de ligação do repressor ao DNA. Quando um dímero de repressor faz contato com seu sítio de ligação ao DNA (Fig. 20.15), as duas regiões 3 da hélice ficam quase inteiramente dentro de sulcos adjacentes maiores em uma face do operador  $\lambda$  do DNA. Cada monômero faz contato com um lado (meio-sítio) da seqüência palindrômica de nucleotídeos em cada sítio de ligação do repressor. As duas regiões 2 das hélices ficam nos sulcos maiores ocupados pela região 3 das hélices.

Como o repressor lambda tem a mais alta afinidade pelos sítios  $O_{L1}$  e  $O_{R1}$ , os primeiros dímeros repressores geralmente se ligam a estes dois sítios. O repressor  $\lambda$  exibe ligação cooperativa com (1) os sítios  $O_{L1}$  e  $O_{L2}$  e (2) os sítios  $O_{R1}$  e  $O_{R2}$ . Assim, a presença de um dímero no sítio  $O_{L1}$  e  $O_{R1}$  aumenta a afinidade do repressor pelos sítios  $O_{L2}$  ou  $O_{R2}$ , respectivamente. Essa cooperatividade não se estende a  $O_{L3}$  ou  $O_{R3}$ . Em concentrações intracelulares normais de repressor, apenas os sítios  $O_{L1}$ ,  $O_{L2}$ ,  $O_{R1}$  e  $O_{R2}$  são ocupados pelo repressor. Quando dímeros de repressor são ligados a  $O_{R1}$ - $O_{R2}$  e  $O_{L1}$ - $O_{L2}$ , a RNA-polimerase não pode ligar-se a  $P_R$  e  $P_L$  e portanto não pode iniciar a transcrição. Assim, os genes  $\lambda$  que codificam funções envolvidas em desenvolvimento lítico são mantidos em um estado reprimido.

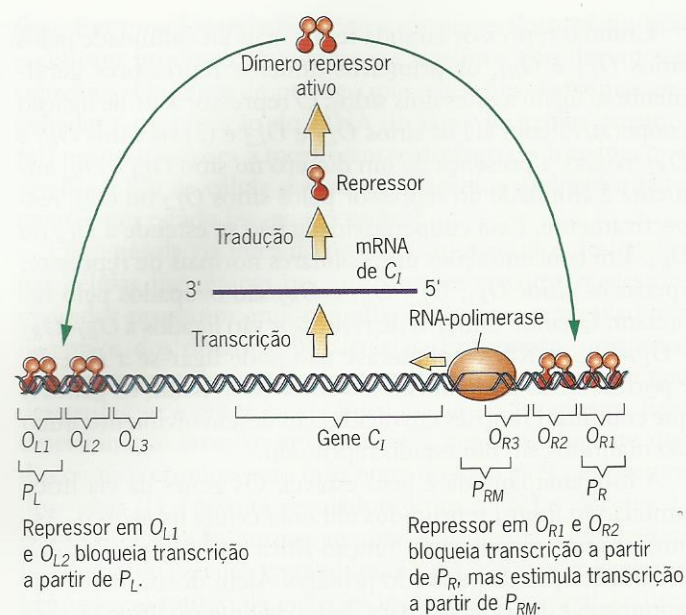
A lisogenia lambda é bem estável. Os genes da via lítica lambda são muito reprimidos em uma célula lisogênica. Assim, produtos gênicos de função lítica não são produzidos por transcrição de genes do profago. Além disso, mudanças espontâneas de lisogenia para desenvolvimento lítico são raras em condições normais de crescimento. Populações de células lisogênicas podem ser induzidas a entrar na via lítica apenas por tratamentos drásticos, tais como irradiação com luz ultravioleta (UV). A irradiação UV de lisógenos lambda ativa uma protease da célula hospedeira que corta a região conectora do repressor  $\lambda$  e a torna não funcional.

Um motivo importante para a estabilidade do estado lisogênico  $\lambda$  é que a síntese do repressor é regulada autogenamente; isto é, a presença do repressor  $\lambda$  estimula a síntese de mais repressor. Por qual mecanismo a síntese do repressor  $\lambda$  é auto-regulada?

O gene  $\lambda C_I$ , que codifica o repressor, é transcrito de um promotor designado  $P_{RM}$  (de promotor para manutenção do repressor), situado entre o gene  $C_I$  e  $O_{R3}$  (Fig. 20.16). O promotor  $P_{RM}$  superpõe-se ao operador  $O_{R3}$ , como o  $P_R$  se superpõe a  $O_{R1}$ . Entretanto, a transcrição iniciada em  $P_{RM}$  ocorre no sentido oposto (para a esquerda, como desenhado na Fig. 20.14) ao da transcrição iniciada em  $P_R$ . O fator fundamental na manutenção do estado lisogênico é que o repressor ligado a  $O_{R2}$  atua como um regulador positivo de transcrição de  $P_{RM}$ . Na ausência do repressor em  $O_{R2}$  (crescimento lítico), a RNA-polimerase não pode ligar-se a  $P_{RM}$  para iniciar a transcrição do gene  $C_I$ . Na presença do repressor, os sítios  $O_{R1}$  e  $O_{R2}$  são ocupados por dímeros do repressor (Fig. 20.16). O repressor ligado a  $O_{R2}$  estimula a ligação da RNA-polimerase a  $P_{RM}$ , resultando na síntese de mais repressor (síntese positiva auto-regulada).

Embora o repressor tenha afinidade mais baixa por  $O_{R3}$  do que por  $O_{R2}$  ou  $O_{R1}$ , ele irá ligar-se a  $O_{R3}$  e bloquear a transcrição a partir de  $P_{RM}$  se presente em alta concentração. Esse bloqueio resultará em concentração mais baixa do repressor, o que, por sua vez, produzirá um óperon  $O_{R3}$  e transcrição renovada de  $P_{RM}$ . O controle da síntese do repressor por este circuito regulador autógeno acentua a estabilidade do estado lisogênico lambda. Ele também ajuda a explicar por que o repressor não é produzido durante o crescimento lítico deste vírus. Se nenhum repressor estiver presente na célula, o sítio



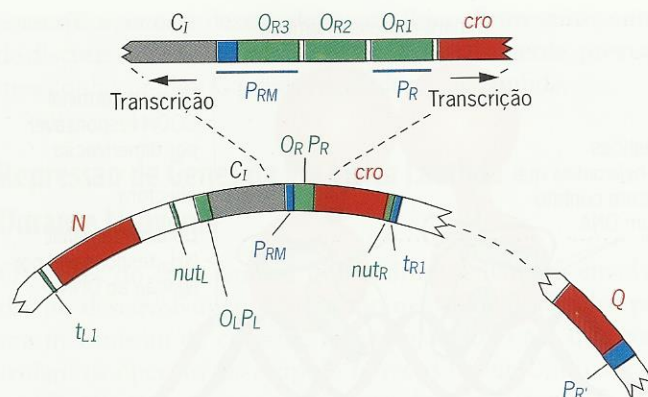


**Fig. 20.16** ■ Regulação autóloga da síntese de repressor do fago lambda. A presença do repressor  $\lambda$  estimula a síntese de mais repressor agindo como regulador positivo de transcrição do gene  $C_I$  a partir do promotor  $P_{RM}$ . Ao mesmo tempo, o repressor funciona como regulador negativo de genes envolvidos no desenvolvimento lítico. Dímeros de repressor ligados a  $O_{L1}$ - $O_{L2}$  e  $O_{R1}$ - $O_{R2}$  impedem a transcrição dos genes de função lítica a partir, respectivamente, dos promotores  $P_L$  e  $P_R$ .

$O_{R2}$  ficará livre, e o repressor não será produzido a partir de transcritos iniciados em  $P_{RM}$ . Como resultado, uma vez que lambda tenha começado a via lítica,  $P_{RM}$  não será usado para produzir repressor, e a ausência de repressor permitirá que o desenvolvimento lítico continue sem interrupção.

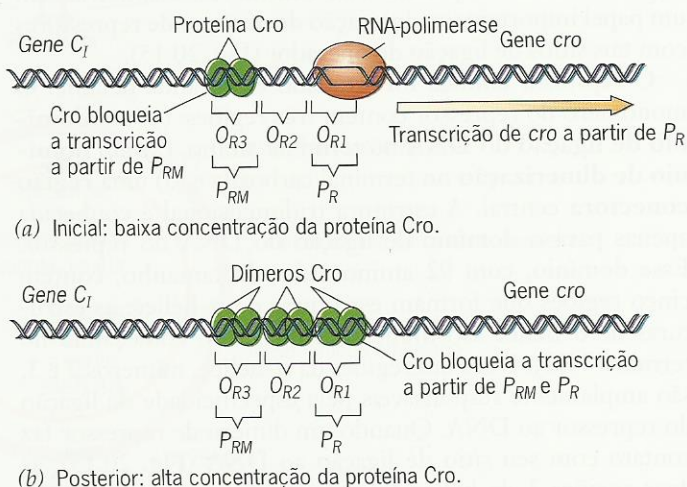
## A Cascata Lítica Reguladora de Lambda

A cascata lítica reguladora de lambda é direta, uma vez que tenha havido a opção de entrar na via lítica. Os genes  $\lambda$  podem ser classificados em três grupos com base em quando eles são expressos durante o ciclo lítico: (1) genes iniciais imediatos, (2) genes iniciais postergados e (3) genes tardios. Lambda tem apenas dois genes iniciais imediatos, *cro* e *N*, ambos codificando proteínas reguladoras necessárias para o desenvolvimento lítico. O cromossomo  $\lambda$  contém uma dúzia de genes iniciais postergados, que especificam produtos necessários para a replicação, recombinação e regulação adicional do DNA. Os 23 genes restantes do fago lambda codificam proteínas com função tardia, envolvidas na morfogênese da cabeça e da cauda e na lise da célula hospedeira. Um dos produtos dos genes iniciais imediatos, a **proteína N**, é necessária para expressar os genes iniciais postergados. Por sua vez, um dos produtos gênicos iniciais postergados, a **proteína Q**, deve estar presente para expressar os genes  $\lambda$  tardios. Vejamos as funções dos genes reguladores fundamentais, *cro*, *N* e *Q*, envolvidos no crescimento lítico do fago  $\lambda$  (Fig. 20.17).



**Fig. 20.17** ■ Genes e sítios de reconhecimento envolvidos na cascata reguladora lítica lambda. Os genes *cro*, *N* e *Q* codificam proteínas reguladoras necessárias para desenvolvimento lítico. Os operadores  $O_L$  e  $O_R$  controlam a transcrição para a esquerda e para a direita a partir dos promotores  $P_L$  e  $P_R$ , respectivamente.  $O_{R1}$ ,  $O_{R2}$  e  $O_{R3}$  são os três sítios de ligação do repressor dentro de  $O_R$ . As sequências *nut\_L* e *nut\_R* são sítios nos quais o produto gênico *N* deve agir para impedir término da transcrição nos sítios  $t_{L1}$  e  $t_{R1}$ , respectivamente.

Na ausência do repressor lambda, o produto do gene  $C_I$  já descrito, a RNA-polimerase inicia transcrição nos promotores  $P_L$  e  $P_R$ . O primeiro gene a ser transcrito começando em  $P_R$  é *cro* (de controle do repressor e outros produtos gênicos). A **proteína cro** é um repressor com uma região de ligação ao DNA muito similar à do repressor  $\lambda$ . Cro liga-se aos mesmos sítios do DNA em  $O_L$  e  $O_R$  que o repressor  $\lambda$ ; mas Cro tem uma afinidade maior por  $O_{R3}$  do que por  $O_{R1}$  e  $O_{R2}$ . Assim, quando Cro é primeiro produzida, ela se liga a  $O_{R3}$  e mantém a síntese do repressor desligada bloqueando o início da transcrição em  $P_{RM}$  (Fig. 20.18a). Posteriormente, quando mais proteína Cro se acumulou na célula, ela tam-



**Fig. 20.18** ■ Funções do produto do gene *cro* nos estágios inicial (a) e tardio (b) de crescimento lítico lambda. A proteína Cro tem afinidade maior por  $O_{R3}$  do que por  $O_{R1}$  ou  $O_{R2}$ , resultando em sua presença apenas em  $O_{R3}$  em baixas concentrações e em todos os três sítios  $O_R$  (e sítios  $O_L$ ) em altas concentrações.



bém se liga a  $O_{R1}$ ,  $O_{R2}$ ,  $O_{L1}$  e  $O_{L2}$ , que suprimem a transcrição a partir de  $P_R$  e  $P_L$ , diminuindo a taxa de síntese dos produtos dos genes iniciais (Fig. 20.18b).

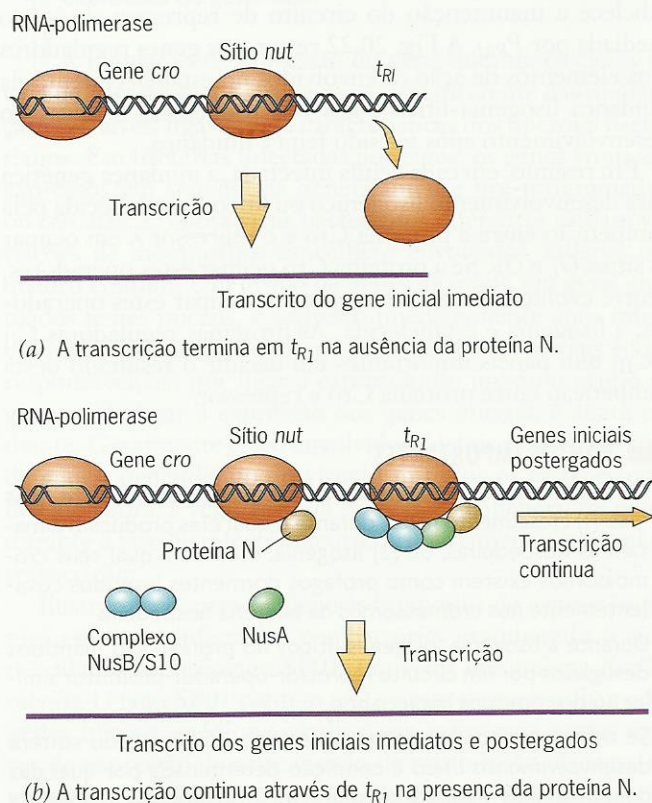
A síntese dos produtos dos genes iniciais postergados requer produtos do segundo gene inicial imediato,  $N$ , que é o primeiro gene transcrito a partir do promotor  $P_L$ . O produto do gene  $N$  funciona para evitar o término da transcrição em dois sítios, um ( $t_{R1}$ ) situado adjacente a  $cro$  e o outro ( $t_{L1}$ ) posterior ao próprio gene  $N$  (Fig. 20.17). Assim, a proteína  $N$  atua como um **antiterminalizador transcricional**. Ela permite que a RNA-polimerase continue a transcrição após os sinais de término  $t_{L1}$  e  $t_{R1}$  e através dos genes iniciais postergados adjacentes. A proteína  $N$  desempenha sua função de antiterminalizador apenas na presença de seqüências de DNA chamadas sítios *nut* (de utilização  $N$ ) situados antecedentes aos sítios terminalizadores (Fig. 20.19). A proteína  $N$  liga-se aos sítios *nut* e, com a ajuda de proteínas de *E. coli* chamadas fatores Nus (de utilização da substância  $N$ ) e proteína ribossômica S10, modifica a especificidade da RNA-polimerase à medida que ela passa do sítio *nut* de modo que não ocorra término quando o complexo RNA-polimerase atinge os sinais de término  $t_{L1}$  e  $t_{R1}$ .

O papel de antiterminalizador da proteína  $N$  é ilustrado na Fig. 20.19. Na ausência da proteína  $N$  funcional, a trans-

crição termina em  $t_{R1}$  e  $t_{L1}$ . Quando presente, a proteína  $N$  liga-se aos sítios *nut* e, juntamente com as proteínas NusA, NusB e S10, modifica a RNA-polimerase de modo que não ocorre término da transcrição em  $t_{R1}$  e  $t_{L1}$ . Como resultado, a transcrição continua nos genes iniciais postergados.

Um dos genes iniciais postergados,  $Q$ , codifica outro antiterminalizador de transcrição que é necessário para expressão gênica tardia. A proteína  $Q$  funciona como a proteína  $N$ , mas reconhece uma seqüência de DNA diferente, chamada *qut* (de utilização de  $Q$ ). A proteína  $Q$  impede a parada da transcrição iniciada no promotor constitutivo  $P_R$ , no sítio de término  $t_{R3}$ , permitindo que a transcrição continue pelos genes tardios.

Profagos lambda não defectivos em *E. coli* lisogênicas podem ser induzidos a entrar na via lítica por exposição a luz ultravioleta (UV). Irradiação das bactérias com UV causa danos ao DNA, o que resulta em várias mudanças fisiológicas em conjunto chamadas de resposta SOS (Cap. 13). Um componente da resposta SOS em *E. coli* é a conversão da proteína RecA (uma importante proteína de recombinação; veja Cap. 13) em uma protease que corta a região conectora do repressor  $\lambda$ . Este evento de clivagem impede a dimerização do domínio de ligação ao DNA do repressor. Na ausência de dímeros ativos de repressor, Cro e proteína  $N$  são produzidas. Cro liga-se a  $O_{R3}$  e impede a síntese do repressor. Com Cro e proteína  $N$ , mas sem repressor, presentes na bactéria, ocorre o desenvolvimento lítico.



**Fig. 20.19** ■ Funções da proteína  $N$  do fago lambda como antiterminalizador de transcrição nos sítios  $t_{R1}$  e (não mostrado)  $t_{L1}$ . (a) Na ausência da proteína  $N$ , a transcrição do gene inicial *cro*, que é iniciada em  $P_R$ , termina em  $t_{R1}$ . (b) Na presença da proteína  $N$ , a transcrição iniciada em  $P_R$  continua através de  $t_{R1}$  nos genes iniciais postergados.

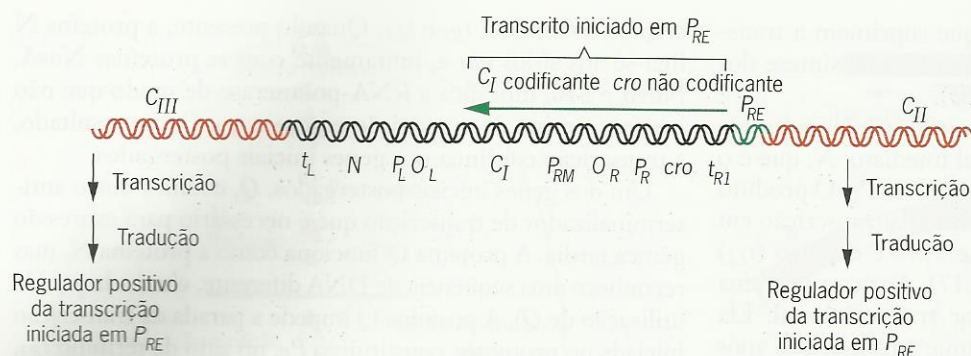
## A Mudança de Lambda: Desenvolvimento Lítico ou Lisogenia

Quando o fago lambda injeta seu DNA na bactéria *E. coli*, esta não contém nem repressor nem proteína Cro. O que então determina se o cromossomo lambda injetado irá (1) entrar na via lisogênica e se inserir no cromossomo hospedeiro ou (2) entrar na via lítica e produzir uma prole de fagos à custa da hospedeira? Tendo examinado os circuitos reguladores lambda envolvidos nas vias lisogênicas e líticas, vamos agora considerar a complexa mudança que controla a ocorrência ou de lisogenia ou de lise após uma infecção de *E. coli* por um fago lambda.

A decisão entre lisogenia e crescimento lítico depende de um equilíbrio delicado entre o repressor  $\lambda$  e a proteína Cro. Que proteína irá ocupar os três sítios de ligação em operadores  $O_L$  e  $O_R$  e, portanto, controlar o início da transcrição em  $P_L$ ,  $P_R$  e  $P_{RM}$ ? A decisão é terminada quando uma destas proteínas ocupa os sítios  $O_L$  e  $O_R$ . Se o repressor  $\lambda$  ocupa estes operadores, resultará lisogenia. Se a proteína Cro ocupa  $O_L$  e  $O_R$ , ocorrerá crescimento lítico.

A mudança genética que direciona lambda para a via lisogênica ou para a via lítica envolve dois genes reguladores principais,  $C_{II}$  e  $C_{III}$ , e um promotor,  $P_{RE}$  (promotor para o estabelecimento do repressor).  $P_{RE}$  está situado do lado oposto do gene *cro* de  $P_{RM}$  (Fig. 20.20). O transcrito iniciado em  $P_{RE}$  contém a seqüência não codificante ou anti-sentido do gene *cro* e a seqüência normal, ou com sentido, do mRNA codificante do gene  $C_I$ . Tanto  $C_{II}$  quanto  $C_{III}$  são genes iniciais





**Fig. 20.20** ■ Elementos reguladores envolvidos em mudança genética lambda para crescimento lisogênico ou para lítico. Os genes  $C_{II}$  e  $C_{III}$  codificam proteínas que são necessárias para início da transcrição no promotor  $P_{RE}$ .

postergados. Portanto, sua expressão é dependente da proteína  $N$ . A proteína  $C_{II}$  é necessária para o início da transcrição em  $P_{RE}$ . A RNA-polimerase pode iniciar a transcrição em  $P_{RE}$  apenas quando a proteína  $C_{II}$  está presente (Fig. 20.21). Entretanto, a proteína  $C_{II}$  é muito instável; ela é rapidamente degradada por uma protease da célula hospedeira. A função da proteína  $C_{III}$  é estabilizar a proteína  $C_{II}$ , permitindo que a RNA-polimerase inicie a transcrição em  $P_{RE}$ . De fato, o efeito de  $C_{III}$  sobre a estabilidade de  $C_{II}$  é uma função das condições de crescimento. Em poucas condições de crescimento,  $C_{III}$  protege  $C_{II}$  de degradação por proteases mais completamente, favorecendo a lisogenia, enquanto em condições ótimas de crescimento,  $C_{III}$  é menos ativa e, portanto,  $C_{II}$  é menos estável, favorecendo o crescimento lítico.

O transcrito iniciado em  $P_{RE}$  favorece o estabelecimento da via lisogênica de dois modos: (1) O repressor lambda é

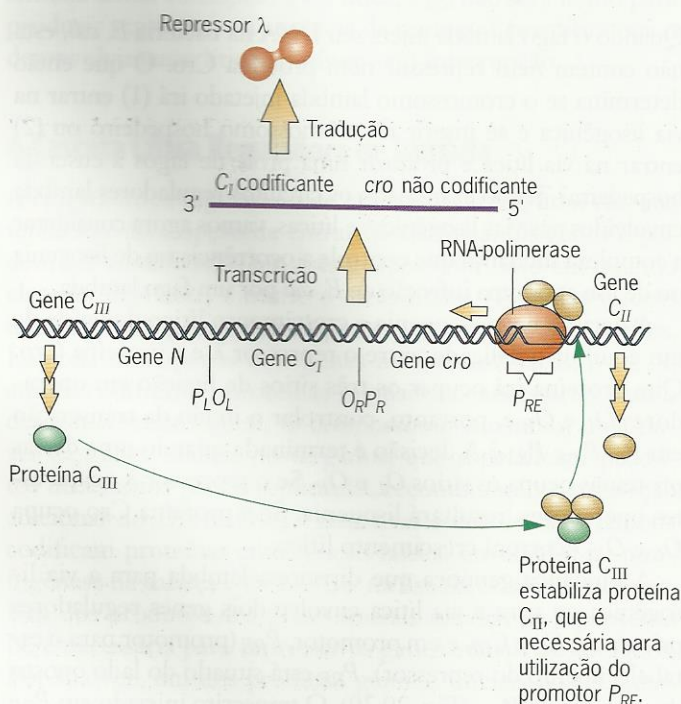
produzido por tradução da sequência codificante  $C_I$ . (2) A parte 5' deste transcrito contém a sequência não codificante ou anti-sentido do gene  $cro$ . Esta sequência anti-sentido  $cro$  é complementar à sequência codificante  $cro$ , ou com sentido, no mRNA  $cro$  autêntico produzido pela transcrição iniciada em  $P_R$ . Assim, esses RNA com sentido e  $cro$  anti-sentido hibridizam-se (formam duplas hélices) uma com a outra, o que bloqueia a tradução do mRNA  $cro$ .

Uma vez que o repressor tenha sido sintetizado pelo uso do circuito  $C_{II}$ - $C_{III}$ - $P_{RE}$ , ele se liga a  $O_{R1}$  e  $O_{R2}$ . A presença do repressor em  $O_{R1}$  e  $O_{R2}$  inibe a expressão dos genes de função lítica bloqueando o início da transcrição em  $P_R$  e estabelece a manutenção do circuito de repressor autógeno mediada por  $P_{RM}$ . A Fig. 20.22 resume os genes reguladores e os elementos de ação *cis* envolvidos no estabelecimento da mudança lisogenia-lítica e nas vias lisogênicas e líticas do desenvolvimento após ter sido feita a mudança.

Em resumo, em cada célula infectada, a mudança genética para desenvolvimento lisogênico ou lítico é estabelecida pela competição entre a proteína Cro e o repressor  $\lambda$  em ocupar os sítios  $O_L$  e  $O_R$ . Se a proteína Cro ocupar estes operadores, ocorre evolução lítica. Se o repressor ocupar estes operadores, a lisogenia é estabelecida. As proteínas reguladoras  $C_{II}$  e  $C_{III}$  têm papéis importantes em decidir o resultado desta competição entre proteína Cro e repressor.

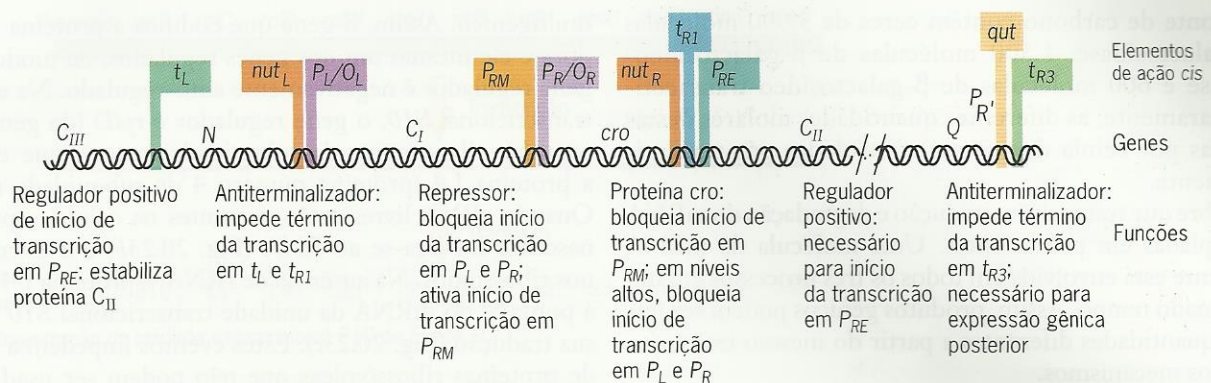
### PONTOS IMPORTANTES

- Os bacteriófagos temperados podem seguir uma dentre duas vias: (1) crescimento lítico, durante o qual eles produzem e matam as hospedeiras, ou (2) lisogenia, durante a qual seus cromossomos existem como profagos dormentes inseridos covalentemente nos cromossomos da bactéria hospedeira.
- Durante a lisogenia, os genes líticos do profago são mantidos desligados por um circuito repressor-operador-promotor similar ao dos óperons bacterianos.
- Se o fago lambda irá entrar no estado lisogênico ou sofrerá desenvolvimento lítico é condição determinada por qual das duas proteínas regulatórias, repressor  $\lambda$  ou proteína Cro, ocupa sítios operadores importantes que controlam a transcrição do genoma  $\lambda$ .
- O estado lisogênico é mantido pelo controle autógeno do repressor  $\lambda$ , enquanto o desenvolvimento lítico é controlado por uma cascata reguladora na qual antiterminalizadores transcricionais exercem um papel principal.



**Fig. 20.21** ■ Diagrama do circuito regulador do fago  $\lambda$   $C_{II}$ - $C_{III}$ - $P_{RE}$  envolvido na determinação sobre se irá ocorrer desenvolvimento lítico ou lisogênico.





**Fig. 20.22** ■ Resumo dos elementos reguladores lambda que controlam desenvolvimento lisogênico e lítico. Elementos de ação *cis* são mostrados em cima, genes codificantes de fatores *trans* são mostrados diretamente acima das moléculas de DNA e as funções dos reguladores de ação *trans* são descritas na parte de baixo. (Baseada em Benjamin Lewin, *Genes V*, Cell Press, 1994. Usada com permissão.)

## SEQUÊNCIAS TEMPORAIS DE EXPRESSÃO GÊNICA DURANTE A INFECÇÃO DE FAGO

Quando o bacteriófago infecta bactérias, o crescimento lítico é controlado por sequências pré-programadas de expressão do gene viral.

A regulação da expressão do gene durante os ciclos de vida de bacteriófagos virulentos é bem diferente das mudanças reversíveis liga-desliga características dos operons bacterianos. Em bactérias infectadas por fagos, os genes virais são expressos em sequências geneticamente pré-programadas, ou cascatas. Embora vírus bacterianos diferentes exibam variações de mecanismos específicos envolvidos, emerge um quadro comum. Um grupo de genes de fagos, em geral chamados genes iniciais, é expresso imediatamente após infecção. O(s) produto(s) de um ou mais dos genes iniciais é(são) responsável(eis) por ligar a expressão do próximo grupo de genes e desligar a expressão dos genes iniciais, e assim em diante. Geralmente estão envolvidos de dois a quatro grupos de genes, dependendo do vírus. Em todos os casos estudados até agora, a regulação da expressão sequencial de genes durante a fase de infecção ocorre primariamente no nível da transcrição.

Ilustraremos estas sequências temporais de expressão gênica em fagos infectados por bactérias examinando o ciclo de vida do bacteriófago SP01, um vírus que infecta *Bacillus subtilis*. O fago SP01 contém três grupos de genes, chamados genes iniciais, médios e tardios, em referência a suas épocas de expressão durante o ciclo de reprodução do fago. Os genes iniciais do fago SP01 são transcritos pela RNA-polimerase de *B. subtilis*. Um dos produtos gênicos iniciais é um polipeptídeo que se liga à RNA-polimerase da célula hospedeira, mudando sua especificidade de modo que a RNA-polimerase modificada transcreve os genes médios de SP01. Dois dos produtos dos genes médios são, por sua vez, polipeptídeos que se associam à RNA-polimerase de *B. subtilis*,

mudando sua especificidade de modo que então transcreva os genes tardios de SP01.

### PONTOS IMPORTANTES

- Sequências temporais pré-programadas de expressão do genoma viral ocorrem em bactérias infectadas por bacteriófagos.
- Os primeiros genes virais expressos em uma bactéria infectada são transcritos pela RNA-polimerase bacteriana não modificada.
- Grupos subsequentes de genes virais expressos são transcritos seja pela RNA-polimerase codificada pelo genoma do fago ou pela RNA-polimerase bacteriana modificada pela adição de proteínas virais.

## CONTROLE TRADUCIONAL DA EXPRESSÃO GÊNICA

A regulação da expressão gênica em geral tem um ajuste fino por modulação seja da frequência de início da tradução seja pela velocidade de alongamento da cadeia polipeptídica.

Embora a expressão gênica em procariontes seja regulada predominantemente no nível da transcrição, o ajuste fino geralmente ocorre no nível da tradução. Em procariontes, as moléculas de mRNA são frequentemente multigênicas, levando sequências codificantes de vários genes. Por exemplo, o mRNA do operon *lac* de *E. coli* abriga sequências de nucleotídeos que codificam  $\beta$ -galactosidase,  $\beta$ -galactosídeo-permease e  $\beta$ -galactosídeo-transacetilase. Assim, os três genes que codificam estas proteínas devem ser ligados e desligados juntos no nível da transcrição porque os genes são co-transcritos. Entretanto, os três produtos gênicos não são feitos em quantidades iguais. Uma *E. coli* que está crescendo em meio rico com lactose como



única fonte de carbono contém cerca de 3.000 moléculas de  $\beta$ -galactosidase, 1.500 moléculas de  $\beta$ -galactosídeo-permease e 600 moléculas de  $\beta$ -galactosídeo-transacetilase. Claramente, as diferentes quantidades molares destas proteínas por célula devem ser controladas pós-transcricionalmente.

Lembre que transcrição, tradução e degradação do mRNA são acopladas em procariontes. Uma molécula de mRNA geralmente está envolvida em todos os três processos em um determinado tempo. Assim, produtos gênicos podem ser feitos em quantidades diferentes a partir do mesmo transcrito por vários mecanismos.

1. Sabe-se que **eficiências desiguais de início de tradução** ocorrem nos códons iniciais ATG de genes diferentes.
2. **Eficiências alteradas de movimento de ribossomos** por regiões intergênicas de um transcrito são bem comuns. Taxas de tradução diminuídas em geral resultam de grampos ou outras formas de estruturas secundárias que impedem a migração do ribossomo ao longo da molécula de mRNA.
3. **Taxas diferenciais de degradação** de regiões específicas de moléculas de mRNA também ocorrem.

A síntese de proteínas ribossômicas de *E. coli* dão vários exemplos bem documentados de regulação translacional da expressão gênica. *E. coli* que estão crescendo rapidamente em condições ótimas precisam de mais ribossomos para a síntese de proteínas do que as que estão crescendo lentamente em condições adversas. Lembre (Cap. 12) que o ribossomo de *E. coli* contém três moléculas de RNA e 52 polipeptídeos. A síntese desses componentes estruturais deve ser coordenada para assegurar estequiometria apropriada para a montagem do ribossomo. Assim, foram desenvolvidos mecanismos reguladores que asseguram a síntese de RNA e proteínas nas quantidades apropriadas, e alguns destes mecanismos reguladores atuam no nível da tradução.

A maioria dos genes de *E. coli* que codificam proteínas ribossômicas estão situados em grupos, e os genes de cada grupo são co-transcritos. Todas exceto uma das proteínas ribossômicas são usadas em quantidades equimolares durante a montagem do ribossomo. Além disso, a síntese das proteínas ribossômicas deve ser coordenada com a síntese dos três RNA ribossômicos. Esta coordenação entre síntese de proteínas ribossômicas e síntese de RNA ribossômico ocorre no nível da tradução.

O grupo de genes da proteína ribossômica S10 (de proteína número 10 da subunidade menor) fornece boa ilustração de como funciona esta regulação. A unidade transcricional S10 contém 11 genes regulados coordenadamente, todos codificando proteínas ribossômicas (Fig. 20.23a). No caso da unidade transcricional S10 e de pelo menos outras cinco destas unidades contendo genes de proteína ribossômica, o produto de um dos genes inibe a tradução do transcrito

multigênico. Assim, o gene que codifica a proteína reguladora é ele mesmo um dos genes regulados, de modo que o gene regulador é negativamente auto-regulado. Na unidade transcricional S10, o gene regulador é *rplD* (de gene *D* da proteína ribossômica da subunidade maior), que codifica a proteína L4 (proteína número 4 da subunidade maior). Quando rRNA livres estão presentes na célula, a proteína nascente L4 liga-se ao RNA (Fig. 20.23b) e é estruturada nos ribossomos. Na ausência de rRNA, a proteína L4 liga-se à ponta 5' do mRNA da unidade transcricional S10 e inibe sua tradução (Fig. 20.23c). Estes eventos impedem a síntese de proteínas ribossômicas que não podem ser usadas pela célula.

A inibição da tradução de uma molécula de mRNA por um dos produtos que ela codifica é comum tanto em procariontes quanto em eucariontes. Este mecanismo é chamado de **auto-regulação negativa** ou **regulação autógena negativa**. Quando o produto gênico é um componente estrutural da célula ou de alguma organela dentro da célula, a auto-regulação em geral é feita pelos monômeros livres presentes na célula.

#### ■ PONTOS IMPORTANTES

- O ajuste fino regulador frequentemente ocorre no nível da tradução por modulação da velocidade de início da cadeia polipeptídica ou alongamento da cadeia.
- Às vezes, a expressão gênica é regulada pela degradação diferencial de regiões específicas de mRNA poligênicos.

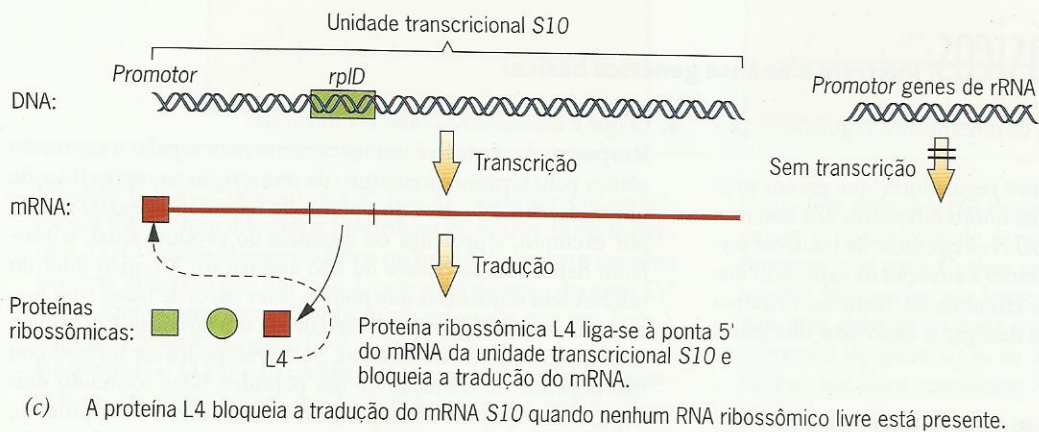
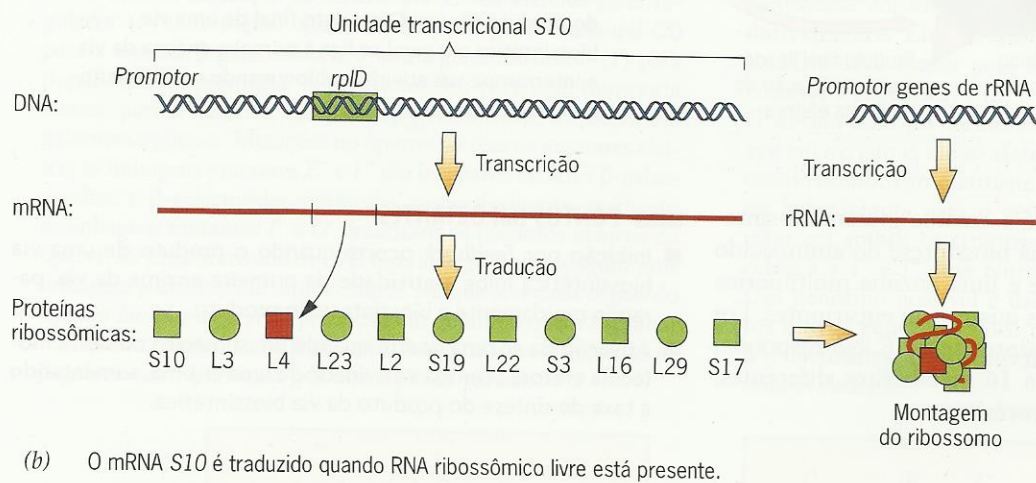
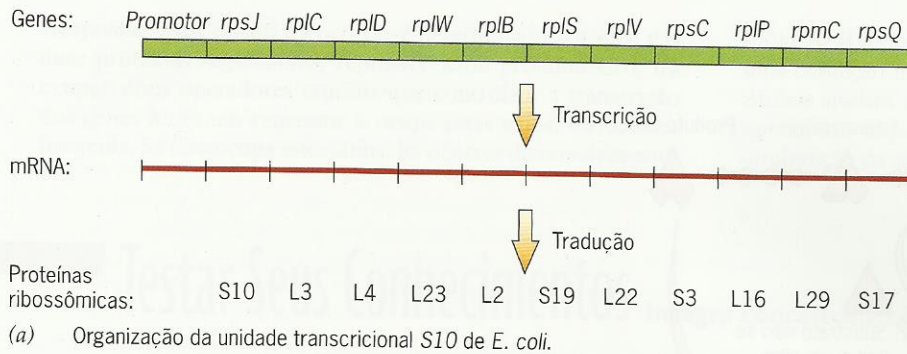
## MECANISMOS REGULADORES PÓS-TRADUCIONAIS

Inibição por *feedback* ocorre quando o produto de uma via biossintética inibe a atividade da primeira enzima da via, desligando rapidamente a síntese do produto.

No início deste capítulo, discutimos o mecanismo pelo qual a transcrição de genes bacterianos codificantes de enzimas em uma via biossintética é reprimido quando o produto da via está presente no meio no qual as bactérias estão crescendo. Um ajuste fino regulador mais rápido do metabolismo em geral ocorre no nível da atividade enzimática. A presença de concentração suficiente do produto final de uma via biossintética frequentemente resulta na inibição da primeira enzima da via (Fig. 20.24). Este fenômeno é chamado **inibição por feedback** ou **inibição pelo produto final**. A inibição por *feedback* resulta em uma parada quase instantânea de síntese do produto final quando este é adicionado ao meio.

A via biossintética de triptofano em *E. coli* fornece uma boa ilustração de inibição por *feedback*. O produto final, triptofano, está ligado à primeira enzima na via, antranilato-sintetase (veja Fig. 20.11), e interrompe completamente sua





**Fig. 20.23** ■ Organização (a) e regulação da tradução (b e c) da unidade transcrricional S10 de *E. coli*, que contém 11 genes codificantes de proteínas ribossômicas. A tradução do mRNA da unidade transcrricional S10 é regulada pela proteína ribossômica L4, que se liga a uma sequência nucleotídica perto da ponta 5' do transcrito S10. (b) Na presença de RNA ribossômico, proteínas ribossômicas interagem com o rRNA na montagem dos ribossomos. (c) Na ausência de RNA ribossômico, a proteína ribossômica L4 liga-se perto da ponta 5' do transcrito S10 e bloqueia sua tradução.

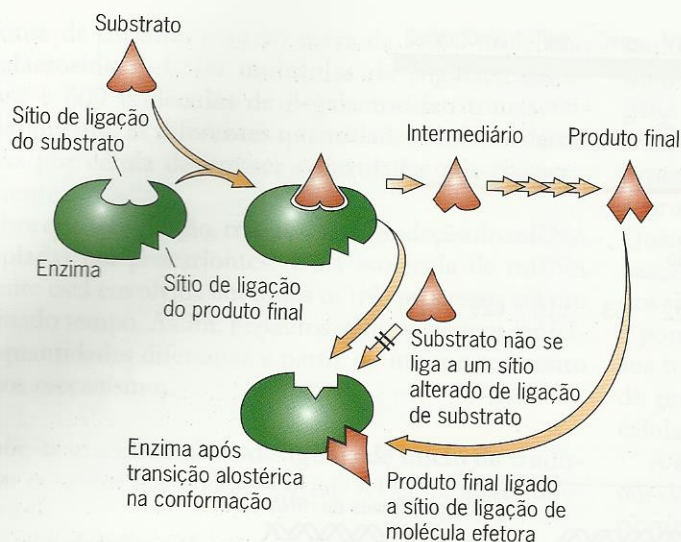
atividade, cessando a síntese de triptofano quase imediatamente.

As enzimas sensíveis a inibição por *feedback* contêm um sítio (ou sítios) de ligação do produto final além do sítio (ou sítios) de ligação do substrato. No caso de enzimas multiméricas, o **sítio de ligação do produto final** ou **regulador** em geral está em uma subunidade (polipeptídeo) diferente do sítio do substrato. Com a ligação do produto final, tais enzimas sofrem transições alostéricas que redu-

zem sua afinidade por seus substratos. As proteínas que sofrem tais mudanças conformacionais são chamadas de proteínas alostéricas. Muitas, talvez a maioria, das enzimas sofrem transições alostéricas de algum tipo.

Transições alostéricas também parecem ser responsáveis por ativação enzimática, que em geral ocorre quando uma enzima se liga a um ou mais de seus substratos ou a alguma outra pequena molécula. Algumas enzimas exibem um amplo espectro de ativação e inibição por muitas moléculas efe-





**Fig. 20.24** ■ Inibição por *feedback* da atividade do produto gênico. O produto final de uma via biossintética em geral se liga à primeira enzima da via e interrompe sua atividade, bloqueando rapidamente a síntese do produto final.

toras diferentes. Um exemplo é a enzima glutamina-sintetase, que catalisa a etapa final na biossíntese do aminoácido glutamina. Glutamina-sintetase é uma enzima multimérica complexa tanto em procariontes quanto em eucariontes. Foi demonstrado que a glutamina-sintetase de *E. coli* responde, seja por ativação ou inibição, a 16 metabólitos diferentes, supostamente por transições alostéricas.

### PONTOS IMPORTANTES

- Inibição por *feedback* ocorre quando o produto de uma via biossintética inibe a atividade da primeira enzima da via, parando rapidamente a biossíntese do produto.
- Ativação da enzima ocorre quando um substrato ou outra molécula efetora acentua a atividade de uma enzima, aumentando a taxa de síntese do produto da via biossintética.

## Exercícios Básicos

Ilustram a análise genética básica.

1. Como podem ser diferenciados os mecanismos reguladores positivos e negativos?  
**Resposta:** As mutações nos genes reguladores que geram produtos não funcionais terão efeitos muito diferentes nos sistemas de controle positivos e negativos. Nos circuitos de controle positivo, tais mutações impossibilitarão a ativação da expressão dos genes regulados, enquanto nos circuitos de controle negativo essas mutações impossibilitarão desligar a expressão dos genes regulados.
2. Como podem ser diferenciados os óperons indutíveis e repressíveis?  
**Resposta:** Na ausência de uma molécula efetora, os óperons indutíveis serão desligados, enquanto os óperons repressíveis serão ligados.
3. Como podem ser diferenciados os elementos reguladores de ação *cis* e *trans*?  
**Resposta:** Eles podem ser distinguidos construindo-se diplóides parciais nos quais os elementos reguladores estão posicionados (1) em *cis* aos genes reguladores e (2) em *trans* aos genes regulados. Um elemento de ação *cis* só influencia a expressão dos genes quando presente na configuração *cis*, enquanto um elemento de ação *trans* exercerá seu efeito na configuração *cis* ou *trans* (compare as Figs. 20.7 e 20.8).
4. O que é atenuação e como ela funciona?  
**Resposta:** Atenuação é um mecanismo para regular a expressão gênica pelo término prematuro da transcrição na região líder de um transcrito. No caso do óperon de triptofano (*trp*) de *E. coli*, por exemplo, a presença ou ausência do produto final, triptofano, determina se ocorre ou não atenuação. A região líder do mRNA tem seqüências que podem fazer pares de bases para formar estruturas de grampios alternativos, um dos quais é um típico sinal de término de transcrição. Se o grampo irá ser formado ou não depende da tradução de um peptídeo líder contendo dois triptofanos. Quando estão presentes baixos níveis de triptofano, a tradução cessa nos códons de Trp, o que impede a formação do grampo de término de transcrição (veja Fig. 20.13b). Quando triptofano suficiente está presente, a tradução continua além dos códons de Trp até o códon de término de tradução, perturbando o primeiro grampo. Isto, por sua vez, permite que seja formado o grampo de término de transcrição e ocorra a atenuação (término da transcrição no atenuador) (veja Fig. 20.13c). A atenuação diminui a síntese das enzimas biossintéticas de triptofano em dez vezes. Atenuação é possível em procariontes porque a transcrição e a tradução estão acopladas, e assim os eventos que ocorrem durante a tradução podem afetar a transcrição.
5. O que determina se o fago lambda irá entrar no estado lítico ou lisogênico quando ele infecta uma *E. coli*?

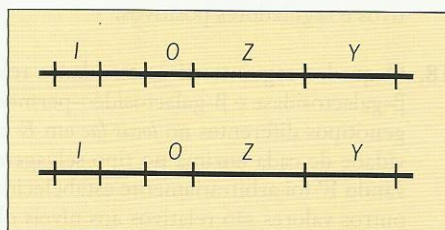


**Resposta:** Qual estado irá ocorrer é determinado por qual das duas proteínas reguladoras, repressor  $\lambda$  ou proteína Cro, irá ocupar sítios operadores cruciais que controlam a transcrição dos genes  $\lambda$ . Se um repressor  $\lambda$  ocupa estes sítios, irá ocorrer lisogenia. Se Cro ocupa estes sítios, irá ocorrer desenvolvimento

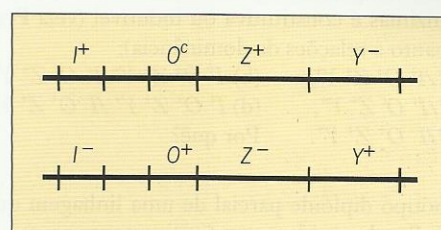
lítico. Se o repressor  $\lambda$  ou Cro irá ligar-se a estes operadores, é uma condição influenciada pelos produtos dos genes  $C_{II}$  e  $C_{III}$ . Ambos ajudam a estabelecer a ligação do repressor  $\lambda$  aos sítios operadores importantes por aumento da síntese do repressor e diminuição da síntese da proteína Cro (veja a Fig. 20.21).

## Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

- O modelo do óperon para a regulação da síntese de enzimas envolvidas na utilização de lactose por *E. coli* inclui o gene regulador (*I*), uma região operadora (*O*), um gene estrutural (*Z*) para a enzima  $\beta$ -galactosidase e outro gene estrutural (*Y*) para  $\beta$ -galactosídeo-permease. A  $\beta$ -galactosídeo-permease transporta lactose para a bactéria, na qual a  $\beta$ -galactosidase a quebra em galactose e glicose. Mutações no óperon *lac* têm os seguintes efeitos: as linhagens mutantes  $Z^-$  e  $Y^-$  são incapazes de fazer  $\beta$ -galactosidase e  $\beta$ -galactosídeo-permease, respectivamente, enquanto as linhagens mutantes  $I^-$  e  $O^c$  sintetizam os produtos gênicos do óperon *lac* constitutivamente. As figuras seguintes mostram uma linhagem parcialmente diplóide de *E. coli* que leva duas cópias do óperon *lac*. No diagrama, preencha um genótipo que irá resultar na síntese constitutiva de  $\beta$ -galactosidase e na síntese indutível de  $\beta$ -galactosídeo-permease por este diplóide parcial.



**Resposta:** Vários genótipos diferentes irão produzir  $\beta$ -galactosidase constitutivamente e  $\beta$ -galactosídeo-permease indutivelmente. Eles devem preencher dois requisitos: (1) a bactéria deve conter pelo menos uma cópia do gene  $I^+$ , que codifica o repressor, e (2) o gene  $Z^+$  bem como uma mutação  $O^c$  devem estar no mesmo cromossomo porque o operador só age em *cis*; isto é, ele só afeta a expressão dos genes no mesmo cromossomo. Em contraste, a bactéria pode ser homozigota ou heterozigota para o gene  $I^+$ , e, se heterozigota,  $I^+$  pode estar em ambos os cromossomos porque  $I^+$  é dominante em relação a  $I^-$  e  $I^+$  atua tanto no arranjo *cis* quanto no *trans*. Um genótipo possível é dado no diagrama seguinte. Quantos outros genótipos você pode imaginar que irão produzir  $\beta$ -galactosidase constitutivamente e  $\beta$ -galactosídeo-permease indutivelmente?



- Quando o fago  $\lambda$  infecta uma *E. coli*, ele pode ter um crescimento lítico e lisar a bactéria hospedeira ou entrar na via lisogênica. (a) O desenvolvimento lítico ou lisogênico ocorre em uma bactéria infectada por um fago  $\lambda$  com uma deleção do gene  $C_I$ ? (b) O que ocorrerá em uma bactéria infectada por um fago  $\lambda$  com uma deleção do gene *Cro*?

**Resposta:** (a) O gene  $C_I$  codifica um repressor  $\lambda$ . Assim, desenvolvimento lítico ocorre nas bactérias infectadas por um fago

$\lambda$  que leva uma deleção do gene  $C_I$ . A proteína Cro irá ocupar os sítios operadores de  $\lambda$  porque nenhum repressor  $\lambda$  estará presente se o gene  $C_I$  tiver sido deletado. (b) Em contraste, lisogenia irá ocorrer nas bactérias infectadas por um fago  $\lambda$  que têm uma deleção do gene *cro*. O repressor será produzido por expressão do gene  $C_I$  e, na ausência da proteína Cro, ele irá ligar-se aos sítios operadores  $\lambda$  e bloquear a expressão dos genes líticos  $\lambda$ .

## Questões e Problemas Acentuam a compreensão e desenvolvem as habilidades analíticas.

- Como podem ser diferenciadas enzimas repressíveis e indutíveis de microrganismos?
- Diferencie entre (a) repressão e (b) inibição por *feedback* causadas pelo produto final de uma via biossintética. Como estes dois fenômenos reguladores complementam um ao outro para proporcionar regulação eficiente do metabolismo?
- No óperon de lactose de *E. coli*, qual a função de cada um dos seguintes genes ou sítios: (a) regulador, (b) operador, (c) promotor, (d) gene estrutural *Z* e (e) gene estrutural *Y*?
- Qual seria o resultado da inativação por mutação dos seguintes genes ou sítios no óperon de lactose de *E. coli*: (a) regulador, (b) operador, (c) promotor, (d) gene estrutural *Z* e (e) gene estrutural *Y*?



- 20.5.** Os seguintes grupos de alelos associados ao óperon de lactose são (em ordem de dominância para cada série alélica): repressor,  $F$  (super-repressor),  $I^+$  (indutível) e  $I^-$  (constitutivo); operador,  $O^c$  (constitutivo, *cis*<sup>-</sup> dominante) e  $O^+$  (indutível, *cis*-dominante); estrutural,  $Z^+$  e  $Y^+$ . (a) Quais dos seguintes genótipos produzirão  $\beta$ -galactosidase e  $\beta$ -galactosídeo-permease se lactose estiver presente: (1)  $I^+ O^+ Z^+ Y^+$ , (2)  $I^- O^c Z^+ Y^+$ , (3)  $F O^c Z^+ Y^+$ , (4)  $F O^+ Z^+ Y^+$  e (5)  $I^- O^+ Z^+ Y^+$ ? (b) Quais destes genótipos produzirão  $\beta$ -galactosidase e  $\beta$ -galactosídeo-permease se lactose estiver ausente? Por quê?
- 20.6.** Por que a mutação  $O^c$  no óperon *lac* de *E. coli* é epistática em relação à mutação  $F$ ?
- 20.7.** Suponha que você descobriu uma nova linhagem de *E. coli* que tem uma mutação na região do operador *lac* que faz com que a proteína repressora tipo selvagem se ligue irreversivelmente ao operador. Você chamou este operador mutante de  $O^{sb}$ , de operador "superbinding". (a) Que fenótipo teria um diplóide parcial com o genótipo  $I^+ O^{sb} Z^- Y^+ / I^+ O^+ Z^+ Y^-$  com relação à síntese das enzimas  $\beta$ -galactosidase e  $\beta$ -galactosídeo-permease? (b) Sua nova mutação  $O^{sb}$  exibe dominância *cis* ou *trans* em seus efeitos sobre a regulação do óperon *lac*?
- 20.8.** Para cada um dos seguintes diplóides parciais, indique se a síntese de enzimas é constitutiva ou indutível (veja Problema 20.5 quanto a relações de dominância):  
 (a)  $I^+ O^+ Z^+ Y^+ / I^+ O^+ Z^+ Y^+$ , (b)  $I^+ O^+ Z^+ Y^+ / I^+ O^c Z^+ Y^+$ ,  
 (c)  $I^+ O^c Z^+ Y^+ / I^+ O^c Z^+ Y^+$ , (d)  $I^+ O^+ Z^+ Y^+ / I^- O^+ Z^+ Y^+$ ,  
 (e)  $I^- O^+ Z^+ Y^+ / I^- O^+ Z^+ Y^+$ . Por quê?
- 20.9.** Escreva o genótipo diplóide parcial de uma linhagem que irá (a) produzir  $\beta$ -galactosidase constitutivamente e permease indutivelmente e (b) produzir  $\beta$ -galactosidase constitutivamente mas não permease seja constitutivamente ou indutivelmente, muito embora o gene  $Y^+$  sabidamente esteja presente.
- 20.10.** Mutações constitutivas produzem níveis enzimáticos elevados em todos os momentos; elas podem ser de dois tipos:  $O^c$  ou  $I^-$ . Suponha que todo o outro DNA presente seja do tipo selvagem. Explique como dois mutantes constitutivos podem ser diferenciados com relação a (a) posição de mapa, (b) regulação dos níveis enzimáticos em diplóides parciais  $O^c/O^+$  versus  $I^-/I^+$  e (c) a posição dos genes estruturais afetados pela mutação  $O^c$  versus os genes afetados por uma mutação  $I^-$  em um diplóide parcial.
- 20.11.** Como um historiador da genética, você está repetindo alguns dos experimentos clássicos conduzidos por Jacob e Monod com o óperon de lactose em *E. coli*. Você usa um plasmídeo  $F'$  para construir várias linhagens de *E. coli* que são parcialmente diplóides para o óperon *lac*. Você constrói linhagens com os seguintes genótipos: (1)  $I^+ O^c Z^+ Y^- / I^+ O^+ Z^- Y^+$ , (2)  $I^+ O^c Z^- Y^+ / I^+ O^+ Z^+ Y^-$ , (3)  $I^- O^+ Z^+ Y^+ / I^+ O^+ Z^- Y^-$ , (4)  $F O^+ Z^- Y^- / I^+ O^+ Z^+ Y^+$  e (5)  $I^+ O^c Z^+ Y^+ / I^+ O^+ Z^- Y^-$ . (a) Qual destas linhagens irá produzir  $\beta$ -galactosidase funcional tanto na presença quanto na ausência de lactose? (b) Qual destas linhagens irá exibir síntese constitutiva de  $\beta$ -galactosídeo-permease funcional? (c) Qual destas linhagens expressará tanto o gene  $Z$  quanto o gene  $Y$  constitutivamente e irá gerar produtos funcionais ( $\beta$ -galactosidase e  $\beta$ -galactosídeo-permease) de ambos os genes? (d) Qual destas linhagens mostrará dominância *cis* de elementos reguladores do óperon *lac*? (e) Qual destas linhagens exibirá dominância *trans* de elementos reguladores do óperon *lac*?
- 20.12.** Como o óperon de triptofano em *E. coli* foi desenvolvido e mantido pela evolução?
- 20.13.** Qual o significado biológico do fenômeno de repressão catabólica?
- 20.14.** Como a concentração de glicose no meio em que foi cultivada a *E. coli* regula o nível intracelular de AMP cíclico?
- 20.15.** O efeito de CAP-cAMP sobre a transcrição do óperon *lac* é um exemplo de regulação positiva ou negativa? Por quê?
- 20.16.** Seria possível isolar mutantes de *E. coli* nos quais a transcrição do óperon *lac* não é sensível a repressão catabólica? Caso sim, em que genes as mutações estariam situadas?
- 20.17.** Usando exemplos, distinga mecanismos reguladores negativos e reguladores positivos.
- 20.18.** O quadro seguinte dá as atividades relativas das enzimas  $\beta$ -galactosidase e  $\beta$ -galactosídeo-permease em células com genótipos diferentes no locus *lac* em *E. coli*. O nível de atividade de cada enzima no tipo selvagem de *E. coli* não levando  $F'$  foi arbitrariamente estabelecido em 100; todos os outros valores são relativos aos níveis observados de atividade nestas bactérias tipo selvagem. Com base nos dados citados no quadro para os genótipos de 1 a 4, preencha os níveis de atividade enzimática que seriam esperados para o quinto genótipo.

| Genótipo                                  | $\beta$ -Galactosidase |           | $\beta$ -Galactosídeo-permease |           |
|-------------------------------------------|------------------------|-----------|--------------------------------|-----------|
|                                           | - Indutor              | + Indutor | - Indutor                      | + Indutor |
| 1. $I^+ O^+ Z^+ Y^+$                      | 0,1                    | 100       | 0,1                            | 100       |
| 2. $I^- O^+ Z^+ Y^+$                      | 100                    | 100       | 100                            | 100       |
| 3. $I^+ O^c Z^+ Y^+$                      | 25                     | 100       | 25                             | 100       |
| 4. $I^- O^+ Z^+ Y^- / F' I^- O^+ Z^+ Y^+$ | 200                    | 200       | 100                            | 100       |
| 5. $I^- O^c Z^- Y^+ / F' I^+ O^+ Z^+ Y^+$ | —                      | —         | —                              | —         |

**20.19.** A taxa de transcrição do óperon *trp* em *E. coli* é controlada por (1) repressão/desrepressão e (2) atenuação. Por que mecanismos estes dois processos reguladores modulam os níveis de transcritos do óperon *trp*?

**20.20.** Que efeito a deleção da região *trpL* do óperon *trp* exerce sobre as taxas de síntese das enzimas codificadas pelos cinco genes no óperon *trp* nas *E. coli* que crescem na presença de triptofano?



- 20.21.** Por que mecanismo a presença de triptofano no meio em que estão sendo cultivadas as *E. coli* resulta em término prematuro ou atenuação da transcrição do operon *trp*?
- 20.22.** Suponha que você usou mutagênese sítio-específica para modificar a sequência de *trpL* de tal modo que dois códons UGG de Trp nas posições 54–56 e 57–60 (veja Fig. 20.12) na sequência líder do mRNA foram modificados para os códons GGG de Gli. A atenuação do operon *trp* ainda será regulada pela presença ou ausência de triptofano no meio em que estão sendo cultivadas as *E. coli*?
- 20.23.** O que é um profago? De que modos o profago do fago  $\lambda$  difere do cromossomo  $\lambda$  presente durante as infecções líticas?
- 20.24.** De que modo a cascata reguladora lítica do fago  $\lambda$  difere da repressão autogenamente mantida de expressão gênica viral que ocorre durante a lisogenia?
- 20.25.** Qual a principal diferença estrutural entre a região do operador no operon *lac* de *E. coli* e os operadores  $O_L$  e  $O_R$  do fago  $\lambda$ ?
- 20.26.** O estado lisogênico  $\lambda$  é muito estável, em parte porque a síntese do repressor  $\lambda$  é regulada autogenamente. O que isto significa? Como a regulação autógena do repressor acentua a estabilidade da lisogenia?
- 20.27.** Em  $\lambda$ , o gene  $C_I$  produz um repressor da via lítica. Mutações  $C_I^-$  causam placas claras, nas quais não estão presentes células lisogênicas. Um plasmídeo possui este alelo mutante  $C_I^-$  sob o controle de seu promotor normal (tipo selvagem). Se este plasmídeo é induzido em uma célula lisogênica que contém um genoma  $\lambda$  tipo selvagem integrado (com um gene  $C_I^+$  normal), o que acontecerá e por quê?
- 20.28.** Durante o crescimento lítico, os genes do fago  $\lambda$  exibem um padrão preciso de expressão temporal. Os genes iniciais imediatos são expressos primeiro, em seguida os genes iniciais postergados e por último os genes tardios. Como funciona esta cascata reguladora? Podem os genes tardios ser expressos sem que os genes iniciais postergados tenham sido primeiro expressos? Caso não, por quê? Que papéis o término da transcrição desempenha na cascata reguladora lítica de  $\lambda$ ?
- 20.29.** A proteína N do fago  $\lambda$  é um antiterminalizador transcripcional. Que papel ela tem na cascata reguladora lambda?
- 20.30.** Um fago lambda tem uma mutação de término de cadeia UAG no meio do gene  $C_I$ . Que fenótipo resultará desta mutação?
- 20.31.** A atenuação do tipo que regula o nível de transcritos de *trp* em *E. coli* provavelmente ocorre em organismos eucarióticos?



# Regulação da Expressão Gênica em Eucariontes

## Tópicos

- Regulação Espacial e Temporal de Genes Eucarióticos
- Modos de Regular a Expressão Gênica Eucariótica
- Indução da Atividade Transcricional por Fatores Ambientais e Biológicos
- Controle Molecular da Transcrição em Eucariontes
- Regulação Pós-transcricional da Expressão Gênica por Interferência de RNA
- Expressão Gênica e Organização Cromossômica
- Ativação e Inativação de Cromossomos Inteiros



Tripanossomas entre hemácias.

## Tripanossomas Africanos: Um Repertório de Disfarces Moleculares

Quase no final do século dezenove, David Bruce, um cirurgião do British Medical Service, resumiu suas observações e experimentos sobre uma doença de animais selvagens e domésticos no sul da África. A doença, chamada nagana, de uma palavra Zulu que significa “perda de espírito”, é caracterizada por febre, tumefação, letargia e emaciação. Bruce reconheceu que nagana é transmitida pela tsé-tsé, um tipo de inseto comum nos espaços abertos da planície africana. Além disso, seu exame de animais doentes o levou a concluir que o agente causal é um protozoário flagelado unicelular que é injetado no sangue do animal durante a picada do inseto tsé-tsé. Este parasita sanguíneo, um tipo



de tripanossoma, hoje é chamado de *Trypanosoma brucei* em homenagem a Bruce. Seres humanos também podem ser infectados por tripanossomas de tsé-tsé, quando então desenvolvem a doença debilitante conhecida como doença africana do sono.

Tanto em animais quanto em seres humanos, as infecções por tripanossomas duram muito tempo. Isto é marcante porque, no sangue, os tripanossomas são submetidos a repetidos ataques pelo sistema imunológico. A cada ataque imunológico, a maioria dos tripanossomas são destruídos. Entretanto, alguns sempre sobrevivem para repopular o sangue e manter a infecção. O segredo deste reaparecimento é a capacidade dos tripanossomas de mudar a capa de proteína de sua superfície. Cada tripanossoma é coberto por cerca de 10 milhões de moléculas de uma única glicoproteína. Quando o sistema imunológico reconhece esta capa de proteína, o tripanossoma infectante sofre ameaça; as células imunológicas o captam e destroem. Entretanto, antes que todos os tripanossomas no animal sejam completamente destruídos, alguns conseguem mudar sua glicoproteína de superfície para uma que não é imediatamente reconhecida pelo sistema imunológico. Esses tripanossomas alterados escapam da destruição e proliferam. Após algum tempo, o sistema imunológico aprende a reconhecê-los também, mas neste meio tempo surge um outro grupo de tripanossomas alterados para manter a continuidade da infecção. Este suprimento aparentemente infindável de disfarces moleculares disponível para os tripanossomas se deve a um grande arranjo de genes que codificam as glicoproteínas variantes da superfície (VSG) que revestem tais organismos. A qualquer momento, apenas um desses genes é expresso; todos os outros estão silenciosos. Entretanto, durante o curso de uma infecção, a identidade do gene expresso é alterada. A cada alteração, os tripanossomas adquirem uma nova proteína de superfície e conseguem ficar um passo adiante das defesas imunológicas do animal. Assim, a infecção é mantida por semanas ou mesmo meses até que, por exaustão, o animal morre.

## REGULAÇÃO ESPACIAL E TEMPORAL DE GENES EUCARIÓTICOS

Os genes eucarióticos são regulados espacialmente e temporalmente.

Os milhares de genes que estão presentes nos genomas de eucariontes multicelulares são normalmente expressos de modo controlado. Uma dimensão deste controle é espacial. Nem todos os produtos gênicos são necessários em todos os tecidos. Alguns genes são expressos em células nervosas, outros em células sangüíneas e ainda outros em células reprodutivas. De fato, a complexidade dos eucariontes multicelulares é parcialmente devida à expressão histo-específica de muitos genes diferentes. Uma segunda dimensão da regulação gênica eucariótica é temporal. Genes diferentes são expressos em momentos diferentes, alguns em resposta a sinais biológicos tais como hormônios e outros em resposta a estímulos ambientais. A especificidade temporal é mais acentuadamente vista durante o desenvolvimento, quando um zigoto se desenvolve em um organismo multicelular. À medida que este organismo se forma, grupos de genes são expressos em uma sequência ordenada para dirigir a forma-

ção de tecidos e órgãos. A regulação temporal e espacial dos genes é, portanto, um aspecto importante dos organismos eucarióticos.

### Regulação Espacial dos Genes de Tubulina em Plantas

Os genes para os polipeptídeos de tubulina são um ótimo exemplo de expressão que é regulada espacialmente. Esses polipeptídeos são os blocos estruturais de microtúbulos (Cap. 2). Existem dois tipos gerais de polipeptídeos de tubulina,  $\alpha$  e  $\beta$ , e uma molécula de cada tipo agrega-se para formar um dímero. Estes dímeros são então dispostos em colunas paralelas para formar microtúbulos cilíndricos ocos com cerca de 24 nm de diâmetro. Vários microtúbulos agregam-se uns aos outros para criar estruturas mais complexas, tais como cílios e flagelos. Dentro das células, os microtúbulos são encontrados em muitos locais – no citoplasma logo abaixo da membrana citoplasmática, ao redor da membrana nuclear e em uma região especial chamada centro de organização de microtúbulos. Os microtúbulos têm um papel importante no movimento celular. Nos cílios e flagelos, seu dobramento ondulante ajuda uma célula a mover-se de uma posição para outra e dentro das células são responsáveis por movimentar os cromossomos durante a mitose.





**Fig. 21.1** ■ Flor de *Arabidopsis thaliana* corada para detectar expressão do promotor do gene de alfa-tubulina *TUA1*. O promotor de *TUA1* foi fundido com a sequência codificante do gene de  $\beta$ -glicuronidase de *Escherichia coli*, que atua como repórter da expressão (cor azul). A expressão de *TUA1* é limitada às anteras.

As tubulinas  $\alpha$  e  $\beta$  são codificadas por conjuntos diferentes de genes. Tipicamente, vários genes diferentes para cada tipo de tubulina estão presentes no genoma. Por exemplo, a planta modelo *Arabidopsis thaliana* tem seis genes de  $\alpha$ -tubulina e nove genes de  $\beta$ -tubulina. Cada um destes genes foi isolado, e a expressão de alguns deles foi estudada em plantas inteiras. Os resultados deste trabalho mostram que cada gene é expresso em um padrão espacial específico. Por exemplo, o gene *TUA1*, que codifica uma variante do polipeptídeo de tubulina  $\alpha$ , é expresso preferencialmente em grãos de pólen (Fig. 21.1). Existe também alguma expressão de *TUA1* em anteras jovens, mas pouca ou nenhuma expressão dele em folhas, caules ou raízes. Em contraste, *TUB1*, um gene de  $\beta$ -tubulina, só é expresso em um trecho da raiz imediatamente acima da ponta da raiz em rápida divisão. Outro gene de  $\beta$ -tubulina, *TUB8*, é preferencialmente expresso nos tecidos vasculares da planta. Os padrões espacialmente específicos de expressão desses e outros genes de tubulina sugerem que tipos ligeiramente diferentes de microtúbulos são necessários em partes diferentes de plantas *Arabidopsis*.

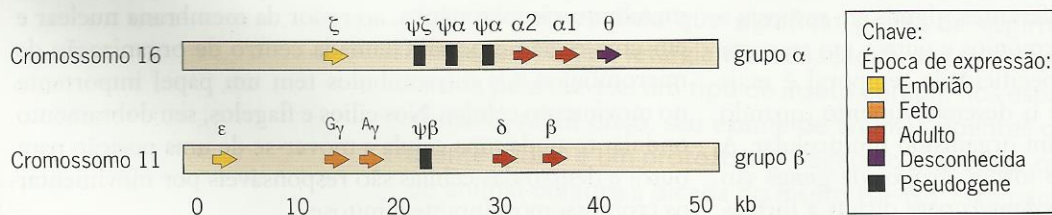
## Regulação Temporal de Genes de Globina em Animais

Talvez um dos exemplos mais marcantes de expressão gênica temporalmente regulada vem do estudo da hemoglobina, a proteína que é responsável por transportar oxigênio no sangue de animais vertebrados. Em vertebrados superiores, esta proteína é um tetrâmero de polipeptídeos chamados globinas. Em cada tetrâmero, existem duas cadeias de globina  $\alpha$  e duas de  $\beta$ . As moléculas de um composto chamado hemo são frouxamente unidas a cada um destes polipeptídeos, formando bolsos que podem ligar oxigênio molecular. Em seres humanos, vários genes para as cadeias de globinas  $\alpha$  e  $\beta$  estão situados em dois sítios separados no genoma. Os genes de globina  $\alpha$  ocupam uma região com 28 kb no cromossomo 16 e os genes de globina  $\beta$  ocupam uma região de 45 kb no cromossomo 11 (Fig. 21.2). Como os genes dentro de cada grupo são duplicados de um gene de globina ancestral, eles formam uma pequena família multigênica. Ao longo do tempo evolutivo, os membros dessas famílias divergiram uns dos outros por mutação aleatória, de modo que hoje em dia cada um codifica um polipeptídeo ligeiramente diferente. Em alguns destes genes duplicados, uma mutação de término de cadeia ou de mudança de matriz de leitura aboliu a capacidade de fazer um polipeptídeo. Tais genes não codificantes são chamados **pseudogenes**. Eles geralmente são indicados pela letra grega  $\psi$  ( $\psi$ ).

Uma característica marcante dos grupos de genes  $\alpha$  e  $\beta$  é que seus membros são expressos em épocas diferentes durante o desenvolvimento. Os genes de um lado nos grupos são expressos apenas no embrião, os do meio são expressos apenas no feto em desenvolvimento e aqueles do outro lado são expressos apenas após o nascimento. Esta ativação sequencial de genes de um lado para outro em um grupo está aparentemente relacionada à necessidade de produzir tipos ligeiramente diferentes de hemoglobina durante o curso do desenvolvimento humano. O embrião, o feto e a criança têm necessidades diferentes de oxigênio, sistemas circulatórios diferentes e ambientes físicos diferentes. A mudança temporal na expressão do gene de globina é aparentemente uma adaptação a esta gama de condições alteradas.

### PONTOS IMPORTANTES

- Os genes de tubulina  $\alpha$  e  $\beta$  em *Arabidopsis* são expressos de modo histoespecífico.
- Os genes de globina  $\alpha$  e  $\beta$  em vertebrados são expressos em épocas diferentes durante o desenvolvimento.





## MODOS DE REGULAR A EXPRESSÃO GÊNICA EUCARIÓTICA

A expressão gênica eucariótica pode ser regulada nos níveis transcricional, de processamento ou de tradução.



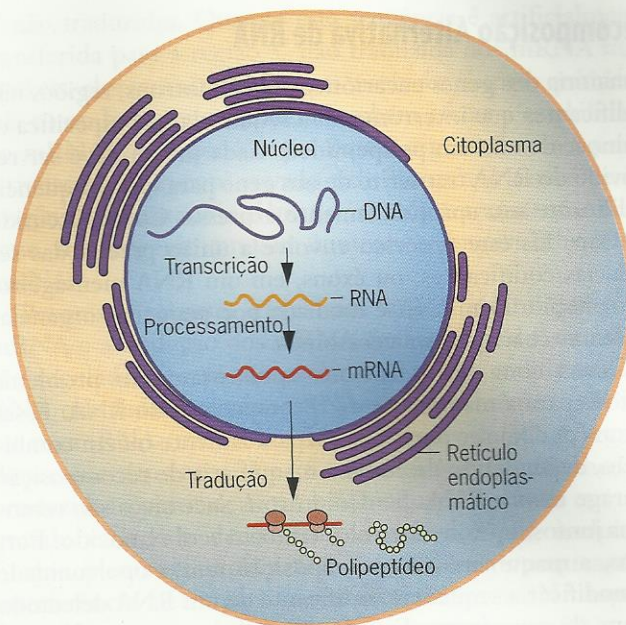
Como em procariontes, a expressão dos genes em eucariontes envolve a transcrição do DNA em RNA e a subsequente tradução deste RNA em polipeptídeos. Mas, antes da tradução, a maioria dos RNA eucarióticos são “processados”. Durante o processamento, o RNA recebe um acréscimo (*cap*) em sua ponta 5', é poliadenilado em sua ponta 3' e é alterado internamente perdendo suas seqüências íntron não codificantes (veja Cap. 11). Os RNA procarióticos tipicamente não sofrem tais modificações terminais e internas.

A expressão gênica em eucariontes é mais complicada do que em procariontes, pois as células eucarióticas são compartimentalizadas por um sistema elaborado de membranas. Essa compartimentalização subdivide as células em organelas separadas, sendo a mais evidente o núcleo. Células eucarióticas também possuem mitocôndrias, cloroplastos (se forem células vegetais) e um retículo endoplasmático. Cada uma dessas organelas desempenha uma função diferente. O núcleo armazena o material genético, as mitocôndrias e os cloroplastos estão associados a energia e o retículo transporta materiais dentro da célula.

A subdivisão de células eucarióticas em organelas separa fisicamente os eventos de expressão gênica. O evento primário, a transcrição do DNA em RNA, ocorre no núcleo. Os RNA transcritos também são modificados no núcleo por adição do *cap*, poliadenilação e remoção de íntrons. Os RNA mensageiros resultantes são então exportados para o citoplasma onde eles se associam aos ribossomos, muitos dos quais estão situados nas membranas do retículo endoplasmático. Uma vez associados aos ribossomos, tais mRNA são traduzidos em polipeptídeos. Essa separação física dos eventos de expressão gênica possibilita que ocorra regulação em diferentes locais (Fig. 21.3). A regulação pode ocorrer no núcleo no nível do DNA ou do RNA ou no citoplasma no nível do RNA ou do polipeptídeo.

### Transcrição Controlada do DNA

Em procariontes, a expressão gênica é regulada principalmente por controle da transcrição do DNA em RNA. Um gene que não é transcrito simplesmente não é expresso. A transcrição ocorre em procariontes quando moléculas reguladoras negativas, tais como a proteína repressora *lac*, foram removidas da vizinhança de um gene e moléculas reguladoras positivas tais como o complexo proteína ativadora do catabolismo (CAP)/AMP cíclico se ligaram a ele (Cap. 20). Estas interações proteína-DNA controlam a possibilidade



**Fig. 21.3** ■ Expressão gênica eucariótica mostrando os estágios nos quais a expressão pode ser regulada: transcrição, processamento e tradução.

de um gene estar ou não acessível à RNA-polimerase. Além disso, os mecanismos que se desenvolveram para controlar a transcrição nestes organismos respondem rapidamente a mudanças ambientais. Como discutimos no Cap. 20, esse controle extremamente sensível é uma estratégia eficiente para a sobrevivência de procariontes.

O controle da transcrição é mais complexo em eucariontes do que em procariontes. Um motivo é que os genes estão seqüestrados no núcleo. Antes que sinais ambientais possam ter qualquer efeito sobre o nível de transcrição, eles devem ser transmitidos desde a superfície da célula, onde são geralmente recebidos, através do citoplasma e da membrana nuclear, até os cromossomos. Células eucarióticas necessitam portanto de sistemas de sinais internos elaborados para controlar a transcrição do DNA. Um outro fator complicador é que muitos eucariontes são multicelulares. Os sinais ambientais têm que passar através de camadas de células para ter um impacto na transcrição dos genes em um determinado tecido. A comunicação intercelular é, portanto, um aspecto importante da regulação transcricional eucariótica.

Como em procariontes, a regulação transcricional eucariótica é mediada por interações proteína-DNA. Proteínas reguladoras positivas e negativas ligam-se a regiões específicas do DNA e estimulam ou inibem a transcrição. Como um grupo, essas proteínas são chamadas de **fatores de transcrição**. Foram identificados muitos tipos diferentes, e a maioria parece ter domínios característicos que lhes permitem interagir com o DNA. A estrutura destas proteínas e a natureza de suas interações com o DNA serão discutidas em uma seção posterior.



## Recomposição Alternativa de RNA

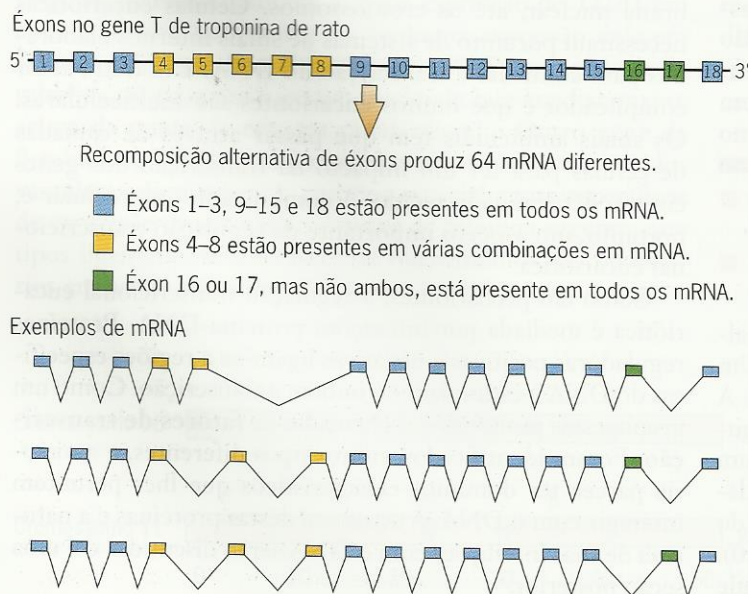
A maioria dos genes eucarióticos possui íntrons, regiões não codificantes que interrompem a sequência que especifica os aminoácidos de um polipeptídeo. Cada íntron deve ser removido do RNA transcrito de um gene para que a sequência codificante seja apropriadamente expressa. Como discutido no Cap. 11, este processo envolve a união precisa das sequências codificantes, ou éxons, em um RNA mensageiro. A formação do mRNA é mediada por pequenas organelas nucleares chamadas spliceossomos.

Genes com vários íntrons apresentam um problema curioso para a maquinaria de recomposição do RNA. Esses íntrons podem ser removidos separadamente ou em combinação, dependendo de como a maquinaria de recomposição interage com o RNA. Se dois íntrons sucessivos são removidos juntos, o éxon entre eles também será removido. Portanto, a maquinaria de recomposição tem a oportunidade de modificar a sequência codificante de um RNA deletando alguns de seus éxons. Este fenômeno de recomposição de um RNA transcrito de modos diferentes aparentemente é um modo de economizar informação genética. Em vez de duplicar os genes, ou trechos de genes, a **recomposição alternativa** dos transcritos possibilita que um único gene codifique polipeptídeos diferentes.

Um exemplo de recomposição alternativa ocorre durante a expressão do gene para troponina T, uma proteína encontrada nos músculos esqueléticos de vertebrados. O tamanho desta proteína varia de cerca de 150 a 250 aminoácidos. No rato, o gene de troponina T tem mais de 16 kb de tamanho e contém 18 éxons diferentes (Fig. 21.4). Os transcritos deste gene são recompostos de modos diferentes para criar uma grande gama de mRNA. Quando eles são traduzidos, muitos polipeptídeos T de troponina diferentes são produzidos. Todos esses polipeptídeos compartilham

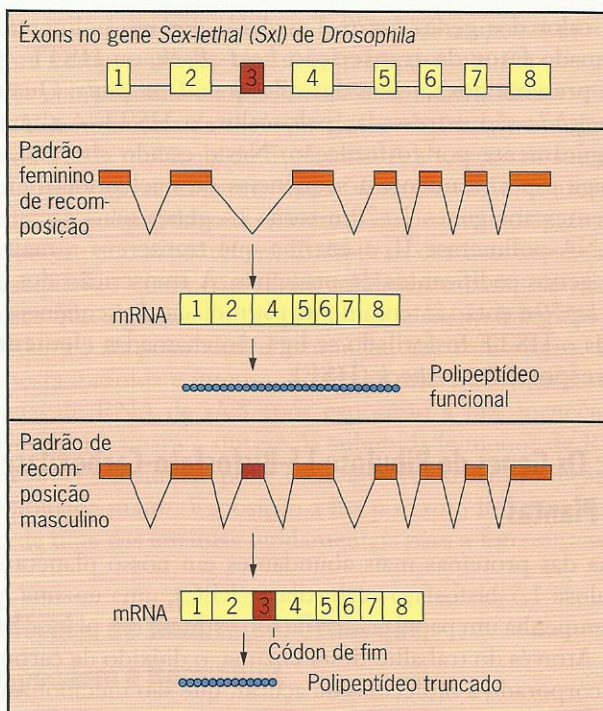
aminoácidos dos éxons 1-3, 9-15 e 18. Entretanto, as regiões codificadas pelos éxons 4-8 podem estar presentes ou ausentes, dependendo do padrão de recomposição, e aparentemente em qualquer combinação. Variação adicional é dada pela presença ou ausência de regiões codificadas pelos éxons 16 e 17; se 16 estiver presente, 17 não estará, e vice-versa. Estas formas diferentes de troponina T supostamente funcionam de modos ligeiramente diferentes dentro dos músculos, contribuindo para a variabilidade da ação muscular.

Outro exemplo de recomposição alternativa ocorre durante a expressão de genes envolvidos na determinação do sexo em *Drosophila*. O regulador principal do processo de determinação do sexo é o gene ligado ao X *Sex-lethal* (*Sxl*). Em fêmeas cromossômicas (XX), o transcrito deste gene é recomposto para produzir um mRNA que codifica uma proteína reguladora. Em machos cromossômicos (XY), o transcrito *Sxl* é alternativamente recomposto para incluir um éxon com um códon de fim. Assim, quando este RNA é traduzido, ele gera um polipeptídeo pequeno sem função reguladora (Fig. 21.5). Em embriões XX, em que a proteína reguladora *Sxl* está presente, um conjunto determinado de genes é expresso fazendo com que tais embriões se desenvolvam como fêmeas. Em embriões XY, em que a proteína reguladora *Sxl* está ausente, é expresso um conjunto diferente de genes que faz com que os embriões se desenvolvam como machos. A recomposição específica de fêmeas ou específica de machos é controlada pela proporção entre cromossomos X e autossomos. Se a proporção X:A é 1,0, ocorre o padrão de recomposição específico de fêmea; se a proporção X:A é 0,5, ocorre o padrão específico de machos. Assim, a recomposição alternativa do RNA *Sxl* é responsável pela diferenciação sexual em *Drosophila*. Maiores detalhes sobre este sistema de determinação do sexo são fornecidos no Cap. 22.



**Fig. 21.4** ■ Recomposição alternativa dos transcritos do gene T de troponina de rato. São mostrados apenas 3 dos 64 possíveis mRNA diferentes.





**Fig. 21.5** ■ Recomposição alternativa dos transcritos do gene *Sex-lethal* em machos e fêmeas de *Drosophila*.

## Controle Citoplasmático da Estabilidade do RNA Mensageiro

Os RNA mensageiros são exportados do núcleo para o citoplasma onde servem como moldes para a síntese de polipeptídeos. Uma vez no citoplasma, um determinado mRNA pode ser traduzido por vários ribossomos que se movem de modo seqüencial. Esta linha de montagem traducional continua até que o mRNA seja degradado. A degradação do RNA mensageiro é portanto outro ponto de controle no processo geral de expressão gênica. Os mRNA de longa vida podem suportar várias rodadas de síntese de polipeptídeos, enquanto os mRNA de vida curta não podem.

Um mRNA que é rapidamente degradado deve ser repostado por transcrição adicional; de outro modo, o polipeptídeo que ele codifica não seria mais produzido. Este término de síntese de polipeptídeos pode, logicamente, ser parte de um programa desenvolvimental. Uma vez que o polipeptídeo exerceu seu efeito, ele pode não ser mais necessário. De fato, sua síntese continuada pode ser prejudicial. Em tais casos, a degradação rápida do mRNA seria um modo razoável de evitar a síntese indesejada do polipeptídeo.

A longevidade do RNA mensageiro pode ser influenciada por vários fatores. As caudas poli(A) parecem estabilizar os mRNA. A seqüência da região 3' não traduzida (3' UTR) que precede uma cauda poli(A) também parece afetar a estabilidade do mRNA. Vários mRNA de vida curta têm a seqüência AUUUA repetida várias vezes em suas regiões

3' não traduzidas. Quando esta seqüência é artificialmente transferida para a região 3' não traduzida dos mRNA mais estáveis, eles também se tornam instáveis. Fatores químicos, tais como hormônios, podem também afetar a estabilidade do mRNA. No sapo *Xenopus laevis*, o gene de *vitelogenina* é transcricionalmente ativado pelo hormônio esteróide estrogênio. Entretanto, além de induzir a transcrição deste gene, o estrogênio também aumenta a longevidade de seu mRNA.

Pesquisas recentes revelaram que a estabilidade dos mRNA e a tradução dos mRNA em polipeptídeos também são reguladas por pequenas moléculas não codificantes de RNA chamadas pequenos RNA de interferência (siRNA) ou microRNA (miRNA). Tais moléculas de RNA reguladoras, que têm entre 21 e 28 nucleotídeos de tamanho, são produzidas a partir de RNA bifilamentares maiores em uma grande variedade de organismos eucarióticos, incluindo fungos, plantas e animais. Os microRNA e os pequenos de interferência fazem par com seqüências em mRNA específicos. Uma vez pareados, eles fazem com que o mRNA seja clivado e subsequentemente degradado, ou impedem que o mRNA seja traduzido em um polipeptídeo. Em plantas, essas pequenas moléculas de RNA fornecem uma defesa crítica contra infecção por vírus com RNA, e tanto em plantas quanto animais elas regulam a expressão de genes envolvidos em maturação e desenvolvimento. Elas serão discutidas em mais detalhes posteriormente neste capítulo.

### PONTOS IMPORTANTES

- Proteínas chamadas fatores de transcrição interagem com o DNA para controlar a transcrição de genes eucarióticos.
- Os transcritos de genes eucarióticos podem ser recompostos alternativamente para produzir RNA mensageiros que codificam polipeptídeos distintos mas correlatos.
- A estabilidade dos RNA mensageiros eucarióticos pode influenciar o nível de síntese de polipeptídeos.

## INDUÇÃO DA ATIVIDADE TRANSCRICIONAL POR FATORES AMBIENTAIS E BIOLÓGICOS

A expressão gênica eucariótica pode ser induzida por fatores ambientais como calor e luz e por moléculas sinalizadoras tais como hormônios e fatores de crescimento.

Em seu estudo do operon de *lactose* em *E. coli*, Jacob e Monod descobriram que os genes para metabolismo de lactose eram transcritos especificamente quando lactose era fornecida às células. Assim, eles demonstraram que a lactose era um **indutor** da transcrição gênica. Seguindo os passos de Jacob e Monod, muitos pesquisadores tentaram identificar indutores específicos da transcrição de genes eucarióticos. Embora tais esforços tenham tido considerável sucesso, a

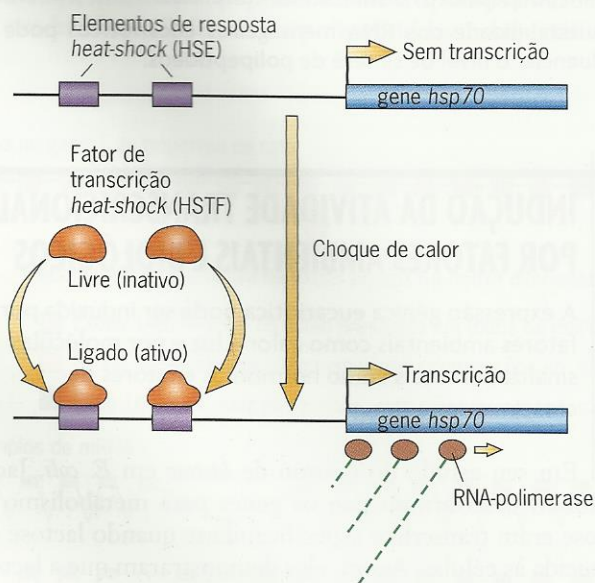


amplitude geral na qual os genes eucarióticos são induzidos por fatores ambientais e nutricionais parece ser menor que a de procariontes. Aqui, consideraremos três exemplos de expressão gênica indutível em eucariotes. As primeiras duas envolvem indução por fatores ambientais – temperatura e luz – e a terceira envolve indução por este grupo especial de moléculas sinalizadoras chamadas hormônios.

### Temperatura: Os Genes *Heat-Shock*

Quando organismos são submetidos a estresse de alta temperatura, eles respondem produzindo um grupo de proteínas que ajudam a estabilizar o ambiente celular interno. Essas **proteínas *heat-shock***, encontradas tanto em procariontes quanto em eucariotes, estão entre os polipeptídeos mais conservados conhecidos. As comparações entre as seqüências de aminoácidos de proteínas *heat-shock* de organismos tão diversos como *E. coli* e *Drosophila* mostram que elas são 40 a 50% idênticas, um achado marcante se considerada a distância de tempo evolutivo que separa estes organismos.

A expressão das proteínas *heat-shock* é regulada em nível transcricional, isto é, o estresse de calor induz especificamente a transcrição dos genes que codificam tais proteínas (Fig. 21.6). Em *Drosophila*, por exemplo, uma das proteínas *heat-shock* chamada **HSP70** (de proteína *heat-shock*, com peso molecular de 70 quilodáltons) é codificada por uma família de genes situados em dois grupos próximos em um dos autossomos. Existem cinco a seis cópias desses genes *hsp70* nos dois grupos. Quando a temperatura excede 33°C, como ocorre nos dias quentes do verão, cada um dos genes é transcrito em RNA, que então é processado e traduzido para produzir polipeptídeos HSP70. Essa transcrição induzida



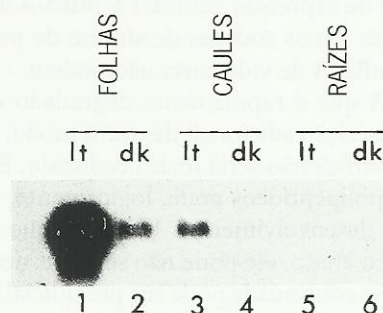
**Fig. 21.6** ■ Indução de transcrição do gene *hsp70* de *Drosophila* por choque provocado por calor. Os HSE estão situados entre 40 e 90 pares de bases antecedentes ao sítio de início da transcrição (seta curva).

por calor dos genes *hsp70* é mediada por um polipeptídeo chamado **fator de transcrição *heat-shock***, ou **HSTF**, que está presente nos núcleos das células de *Drosophila*. Quando *Drosophila* são estressadas pelo calor, o HSTF é alterado quimicamente por fosforilação. Neste estado alterado, ele se liga especificamente a seqüências de nucleotídeos antecedentes aos genes *hsp70* e torna os genes mais acessíveis à RNA-polimerase II, a enzima que transcreve a maioria dos genes codificantes de proteínas. A transcrição dos genes *hsp70* é então intensamente estimulada. As seqüências às quais o HSTF fosforilado se liga são chamadas **elementos de resposta *heat-shock*** (HSE).

### Luz: Os Genes de Ribulose 1,5-Bisfosfato-Carboxilase em Plantas

Uma das proteínas mais abundantes em nosso planeta é a ribulose 1,5-bisfosfato-carboxilase (RBC), uma enzima que desempenha um papel crítico na fotossíntese das plantas verdes. Através do trabalho desta enzima, o dióxido de carbono é incorporado a moléculas de açúcar, que são metabolizadas para fornecer energia para as células. Este processo depende da capacidade das plantas de absorver energia luminosa. Sem luz, todo o processo é interrompido, e não há necessidade de enzimas tais como RBC. Portanto, não é surpresa que a produção de RBC seja induzida especificamente quando plantas são expostas à luz.

A RBC é uma enzima complexa que consiste em subunidades grandes e pequenas, cada uma codificada por genes diferentes. Em algumas espécies de plantas, ambos os genes estão situados no DNA do cloroplasto (Cap. 19), mas em outras o gene para a subunidade menor está situado no DNA nuclear e o gene para a subunidade maior está situado no DNA do cloroplasto. A expressão do gene nuclear para a subunidade menor (indicado por *rbcS*) foi analisada em diferentes espécies de plantas. Tais estudos mostraram que *rbcS* é intensamente transcrito após exposição à luz (Fig. 21.7).



**Fig. 21.7** ■ Transcrição induzida por luz do gene para a subunidade menor de ribulose 1,5-bisfosfato-carboxilase. As plantas foram ou expostas à luz (Lt) ou mantidas no escuro (dk) e o RNA foi extraído de diferentes tecidos da planta. O DNA extraído foi analisado por transferência northern usando-se uma sonda marcada radioativamente para o gene *rbcS*. O auto-radiograma resultante mostra que o RNA de *rbcS* é produzido em folhas e caules, especialmente após exposição à luz.



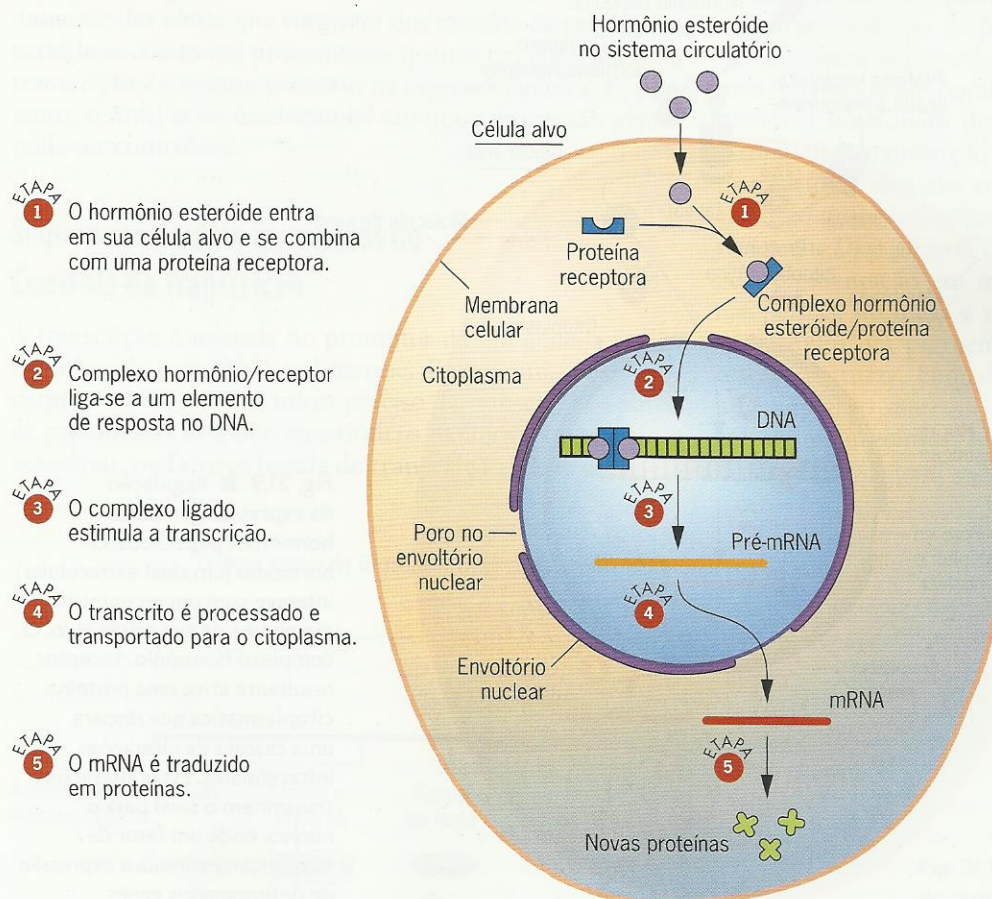
A transcrição induzida por luz de *rbcS* é apenas parcialmente compreendida. Um evento importante é a absorção de luz pela proteína citoplasmática chamada fitocromo (Cap. 19). Uma molécula de absorção de luz chamada cromóforo é ligada a cada molécula de fitocromo. A absorção de luz pelo cromóforo aparentemente causa uma alteração conformacional no polipeptídeo do fitocromo, que então dispara mudanças em outras proteínas. Embora os detalhes dos eventos subsequentes ainda não sejam totalmente compreendidos, parece que algumas destas proteínas se ligam a regiões antecedentes ao gene *rbcS* e estimulam sua transcrição. Esta resposta é rápida e intensa. É gerada uma quantidade substancial de RNA de *rbcS* após exposição a luz, fornecendo vários moldes para a síntese da subunidade menor de RBC. Um processo similar pode estar envolvido na produção da subunidade maior. Exposição a luz portanto induz a produção de uma das enzimas fundamentais para a fotossíntese.

## Moléculas de Sinalização: Genes que Respondem a Hormônios

Em eucariontes multicelulares, um tipo de célula pode ativar outra secretando um **hormônio**. Os hormônios circulam pelo corpo, fazem contato com suas células alvos e então iniciam uma série de eventos que regulam a expressão de

determinados genes. Em animais, existem duas classes gerais de hormônios. A primeira classe, os **hormônios esteróides**, são moléculas pequenas lipossolúveis derivadas do colesterol. Devido a sua natureza lipídica, elas têm pouca ou nenhuma dificuldade em atravessar as membranas celulares. Os exemplos são estrogênio e progesterona, que têm papéis importantes nos ciclos reprodutivos femininos, testosterona, um hormônio de diferenciação e comportamento masculino, os glicocorticóides, que estão envolvidos na regulação dos níveis de açúcar no sangue, e ecdisona, um hormônio que controla os eventos do desenvolvimento em insetos. Uma vez que estes hormônios entram na célula, eles interagem com proteínas citoplasmáticas ou nucleares chamadas **receptores hormonais**. O complexo hormônio/receptor que é formado então interage com o DNA onde ele atua como um fator de transcrição para regular a expressão de alguns genes (Fig. 21.8).

A segunda classe de hormônios, os **hormônios peptídicos**, são cadeias lineares de aminoácidos. Como todos os outros polipeptídeos, estas moléculas são codificadas por genes. Os exemplos são insulina, que regula os níveis de açúcar sanguíneo, somatotrofina, que é um hormônio de crescimento, e prolactina, que tem como alvo tecidos nas mamas de fêmeas de mamíferos. Como os hormônios peptídicos são tipicamente muito grandes para passar livremente pelas membranas celulares, os sinais que eles levam devem



**Fig. 21.8** ■ Regulação da expressão gênica por hormônios esteróides. O hormônio interage com um receptor dentro de sua célula alvo. Neste exemplo, o receptor está no citoplasma; outros receptores de hormônios esteróides estão situados no núcleo. Neste exemplo, o complexo esteróide/receptor de hormônio move-se para o núcleo, onde ativa a transcrição de genes particulares.

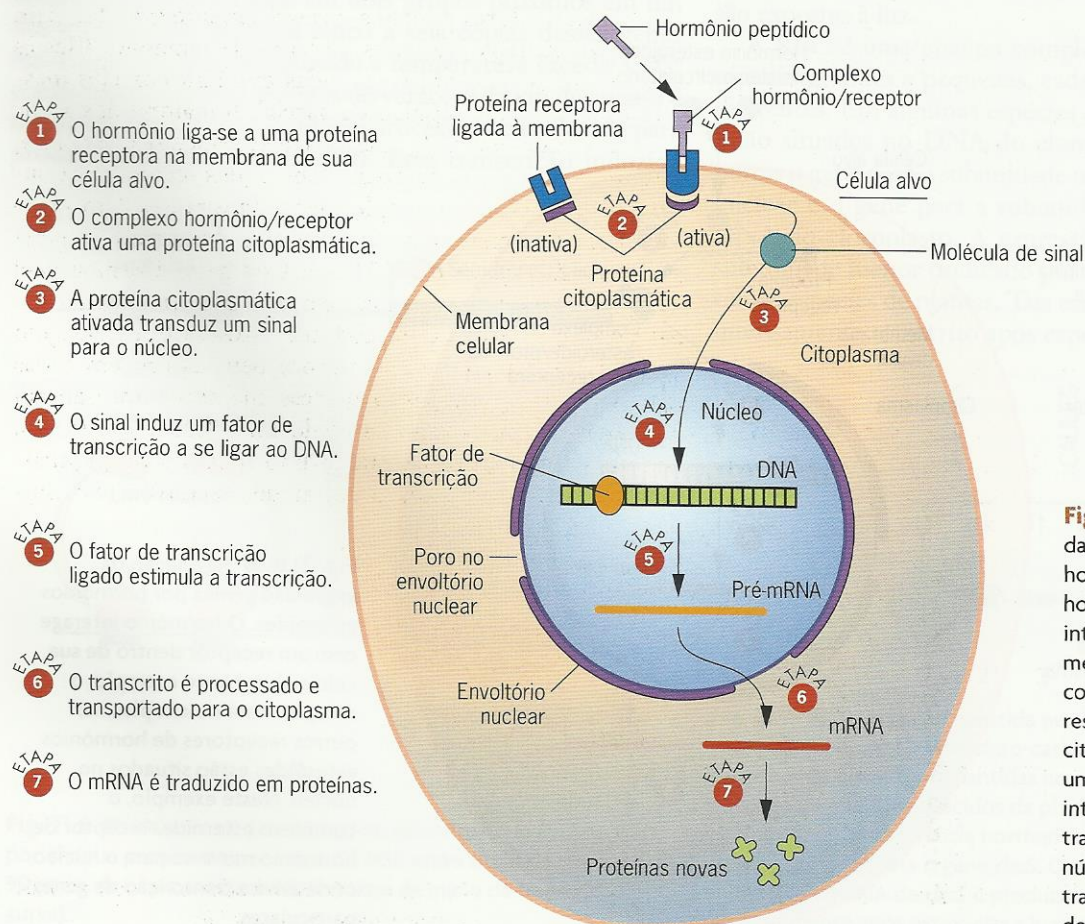


ser transmitidos para o interior das células por **proteínas receptoras ligadas à membrana** (Fig. 21.9). Quando um hormônio peptídico interage com seu receptor, ele causa uma mudança conformacional no receptor que acaba por resultar em mudanças em outras proteínas dentro da célula. Por meio de uma cascata de tais mudanças, o sinal hormonal é transmitido através do citoplasma da célula para o núcleo, onde finalmente tem o efeito de regular a expressão de genes específicos. Este processo de transmissão de sinais hormonais através da célula em direção ao núcleo é chamado de **transdução de sinal**.

A expressão gênica induzida por hormônios é mediada por seqüências específicas no DNA. Essas seqüências, chamadas **elementos de resposta hormonal (HRE)**, são análogas aos elementos de resposta *heat-shock* já discutidos. Estão situadas perto de genes que elas regulam e servem para a ligação de proteínas específicas, que então atuam como fatores de transcrição. Com hormônios esteróides tais como o estrogênio, os HRE são ligados pelo complexo hormônio/receptor, que então estimula a transcrição. O vigor desta resposta transcricional depende do número de HRE presentes. Quando existem vários elementos de resposta, complexos hormônio/receptor ligam-se cooperativamente, aumentando de maneira significativa a taxa de transcrição, isto é, um gene com dois elementos de resposta

é transcrito duas vezes mais intensamente do que um gene com apenas um. Com hormônios peptídicos, o receptor geralmente permanece na membrana celular, mesmo após ter formado um complexo com o hormônio. O sinal hormonal é portanto levado para o núcleo por outras proteínas, algumas das quais se ligam a seqüências perto dos genes que são regulados pelo hormônio. Estas proteínas então atuam como fatores de transcrição para controlar a expressão dos genes.

A atividade transcricional pode ser induzida por muitos outros tipos de proteínas que não são hormônios no sentido clássico, isto é, não são produzidos por uma determinada glândula ou órgão. Eles incluem uma variedade de moléculas circulantes secretadas tais como fator de crescimento nervoso, fator de crescimento epidérmico, fator de crescimento derivado de plaquetas e outras moléculas não circulantes associadas a superfícies celulares ou à matriz entre as células. Embora cada uma destas proteínas tenha suas próprias peculiaridades, o mecanismo geral pelo qual induzem transcrição assemelha-se ao dos hormônios peptídicos. Uma interação entre a proteína sinalizadora e um receptor ligado a membrana inicia uma cadeia de eventos dentro da célula que no final resultam em fatores específicos de transcrição que se ligam a determinados genes, que são então transcritos.



**Fig. 21.9** ■ Regulação da expressão gênica por hormônios peptídicos. O hormônio (um sinal extracelular) interage com um receptor na membrana de sua célula alvo. O complexo hormônio/receptor resultante ativa uma proteína citoplasmática que dispara uma cascata de alterações intracelulares. Essas alterações transmitem o sinal para o núcleo, onde um fator de transcrição estimula a expressão de determinados genes.



### PONTOS IMPORTANTES

- A transcrição dos genes *hsp70* em resposta a temperatura aumentada é mediada por um fator de transcrição *heat-shock*.
- A transcrição do gene para a enzima fotossintética ribulose 1,5-bisfosfato-carboxilase (RBC) é induzida por exposição à luz.
- Hormônios esteróides e suas proteínas receptoras formam complexos que atuam como fatores de transcrição para regular a expressão de genes específicos.
- Hormônios peptídicos interagem com proteínas receptoras ligadas a membrana para ativar um sistema de sinalização que regula a expressão de genes específicos.

## CONTROLE MOLECULAR DA TRANSCRIÇÃO EM EUKARIOTES

A transcrição de genes eucarióticos é regulada por interações entre proteínas e seqüências de DNA dentro ou perto dos genes.

Grande parte das pesquisas atuais sobre expressão gênica eucariótica enfoca os fatores que controlam a transcrição. Esta intensa ênfase no controle transcricional parcialmente se deve ao desenvolvimento de técnicas experimentais que permitiram o estudo em grande detalhe deste aspecto da regulação gênica. Entretanto, também é decorrente do destaque das idéias que surgiram dos estudos de genes procarionóticos. Tanto em procarionotes quanto em eucariotes, a transcrição é o evento primário na expressão gênica. É, portanto, o nível mais fundamental em que a expressão gênica pode ser controlada.

### Seqüências de DNA Envolvidas no Controle da Transcrição

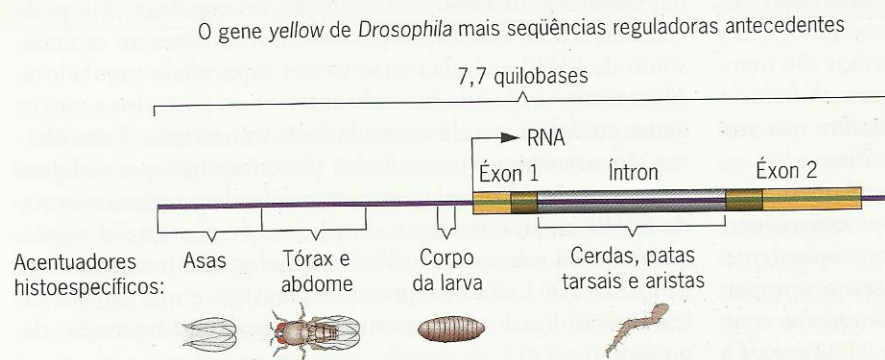
A transcrição é iniciada no promotor de um gene, a região reconhecida pela RNA-polimerase. Entretanto, como discutimos no Cap. 11, o início preciso da transcrição a partir de promotores de genes eucarióticos requer várias proteínas acessórias, ou **fatores basais de transcrição**. Cada uma des-

as proteínas se liga a uma seqüência dentro do promotor para facilitar o alinhamento preciso da RNA-polimerase no filamento molde do DNA.

A transcrição dos genes eucarióticos também é controlada por uma variedade de **fatores especiais de transcrição**, tais como os envolvidos na regulação de genes induzidos por calor, luz e hormônios que já discutimos. Estes fatores ligam-se a elementos de resposta ou, mais geralmente, a seqüências chamadas de **acentuadores** situadas na vizinhança do gene. Os fatores especiais de transcrição que se ligam a estes acentuadores podem interagir com fatores basais de transcrição e com a RNA-polimerase, que se liga ao promotor de um gene. As interações que ocorrem entre os fatores especiais de transcrição, os fatores basais de transcrição e a RNA-polimerase regulam a atividade transcricional de um gene.

Os acentuadores exibem três propriedades mais ou menos gerais: (1) eles atuam em distâncias relativamente grandes – de até vários milhares de pares de bases de seus genes regulados; (2) sua influência na expressão gênica é independente de orientação – eles funcionam igualmente bem seja na orientação normal ou invertida dentro do DNA; e (3) seus efeitos são independentes de posição – eles podem estar situados antecedentes ou posteriores a um íntron de um gene ou dentro dele e ainda exercer grande efeito sobre a expressão do gene. Estas três características distinguem os acentuadores dos promotores, que estão situados tipicamente logo antecedendo o gene e funcionam apenas em uma orientação.

Os acentuadores podem ser relativamente grandes, com até várias centenas de pares de bases de tamanho. Eles às vezes contêm seqüências repetidas que têm por eles mesmos uma atividade reguladora parcial. A maioria dos acentuadores funcionam de maneira histoespecífica, ou seja, estimulam a transcrição apenas em alguns tecidos. Em outros tecidos, eles são simplesmente ignorados. Um claro exemplo desta histoespecificidade vem do estudo do gene *yellow* em *Drosophila* (Fig. 21.10). Este gene é responsável pela pigmentação em muitas partes do corpo – nas asas, nas pernas, no tórax e no abdome. Moscas tipo selvagem apresentam um pigmento marrom bem escuro em todas estas estruturas, enquanto moscas mutantes apresentam um pigmento marrom-amarelado mais claro. Entretanto, em alguns mutantes, há um padrão mosaico de pigmentação, marrom bem escuro em alguns tecidos e marrom-



**Fig. 21.10** ■ Os acentuadores histoespecíficos do gene *yellow* de *Drosophila*.





## ENFOQUE TÉCNICO: GAL4, um Fator de Transcrição de Leveduras

Células de leveduras cultivadas em um meio que contenha galactose mas não glicose expressam especificamente vários genes cujos produtos estão envolvidos no metabolismo de glicose. Dois destes genes, *GAL1* e *GAL10*, estão separados por menos de uma quilobase nos cromossomos de leveduras (Fig. 1). A transcrição destes genes é controlada por uma sequência situada entre eles. Esta sequência, chamada *seqüência ativadora antecedente*, ou *UAS*, é o sítio de ligação para uma proteína que ativa a transcrição a partir dos promotores *GAL1* e *GAL10*. Como tais promotores estão orientados em sentidos opostos, os genes de *GAL1* e *GAL10* são transcritos divergentemente.

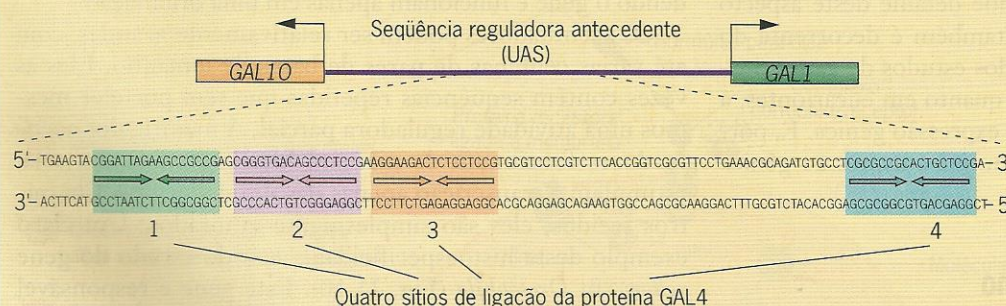
A proteína que se liga a UAS entre *GAL1* e *GAL10* é codificada por um gene não ligado, *GAL4*. Este gene é constitutivamente transcrito em células de leveduras. Na presença de galactose, a proteína *GAL4* é capaz de ativar a transcrição de *GAL1* e *GAL10*, mas, na ausência de galactose, a proteína *GAL4* não pode ativar a transcrição destes genes. Assim, a galactose pode ser considerada um indutor da expressão gênica de *GAL1* e *GAL10* – por um efeito da proteína *GAL4*.

A proteína *GAL4* tem 881 aminoácidos de tamanho. Experimentos com proteínas *GAL4* truncadas mostraram que os primei-

ros 73 aminoácidos são suficientes para ligação à UAS. Esta parte da proteína contém um motivo *zinc finger*, uma característica estrutural frequentemente associada à capacidade de se ligar ao DNA. Uma vez que a proteína *GAL4* se tenha ligado ao DNA, ela pode ativar a transcrição de genes vizinhos. Estudos com partes da proteína *GAL4* mostraram que esta propriedade ativadora está em dois domínios separados. Um vai dos aminoácidos 148 a 196 e o outro dos aminoácidos 768 a 881. A ativação transcricional parece envolver o contato entre estas partes da proteína *GAL4* e outras proteínas que se ligaram ao promotor do gene, incluindo, possivelmente, a RNA-polimerase II.

As UAS situadas entre os genes *GAL1* e *GAL10* assemelham-se aos acentuadores encontrados em eucariotes superiores. Elas funcionam em ambas as orientações e a alguma distância dos genes que regulam. Entretanto, ao contrário do acentuador, elas não podem funcionar posteriores ao sítio de início de transcrição do gene. Análise molecular revelou que as UAS contêm quatro sítios de ligação de *GAL4*. Cada um é um palíndromo imperfeito de oito pares de bases centrado em um par de bases A:T.

Quando a UAS é artificialmente inserida antecedendo um gene, este gene fica sob o controle da proteína *GAL4*. Este fe-



**Fig. 1** ■ Os genes *GAL1* e *GAL10* em leveduras. A proteína *GAL4* regula de maneira coordenada estes genes divergentemente transcritos ligando-se a quatro sítios dentro da seqüência ativadora antecedente (UAS) entre eles.

amarelado em outros. Estes padrões em mosaico se devem a mutações que alteram a transcrição do gene *yellow* em alguns tecidos mas não em outros. Pamela Geyer e Victor Corces mostraram que o gene *yellow* é regulado por vários acentuadores, alguns dos quais estão situados dentro de um íntron, e que cada acentuador ativa a transcrição em um tecido diferente. Se, por exemplo, o acentuador para expressão na asa está mutado, as nervuras nas asas são marrom-amareladas em vez de marrom bem escuro. A bateria de acentuadores associados ao gene *yellow* permite que sua expressão seja controlada de modo histoespecífico.

Um dos primeiros acentuadores a serem amplamente estudados está situado no cromossomo do vírus eucariótico SV40. Este vírus infecta células de macaco e foi amplamente usado em pesquisas biológicas. Seu cromossomo circular de 5,2 kb contém um único acentuador proeminente com cerca de 220 pares de bases de tamanho (Fig. 21.11a, b). O

acentuador regula a transcrição de dois grupos de genes no cromossomo do vírus. Um grupo, situado à direita do acentuador, é transcrito bem cedo durante a infecção e o outro grupo, situado à esquerda, é transcrito depois. O acentuador SV40 contém duas repetições, diretas com 72 pb, cada uma das quais é suficiente para a função acentuadora. Ele pode estar invertido ou movido para locais diferentes no cromossomo de SV40 e ainda conserva sua capacidade reguladora. Além disso, se ele está inserido anterior ou posterior a outros genes eucarióticos, ele estimula a sua transcrição. Estes efeitos são supostamente mediados por proteínas que se ligam ao acentuador. Curiosamente, examinando os cromossomos de SV40 ao microscópio eletrônico, vemos que a região acentuadora não está enrolada ao redor dos nucleossomos (Fig. 21.11c). Uma interpretação plausível é que fatores de transcrição ligados ao acentuador impedem a formação do nucleossomo.





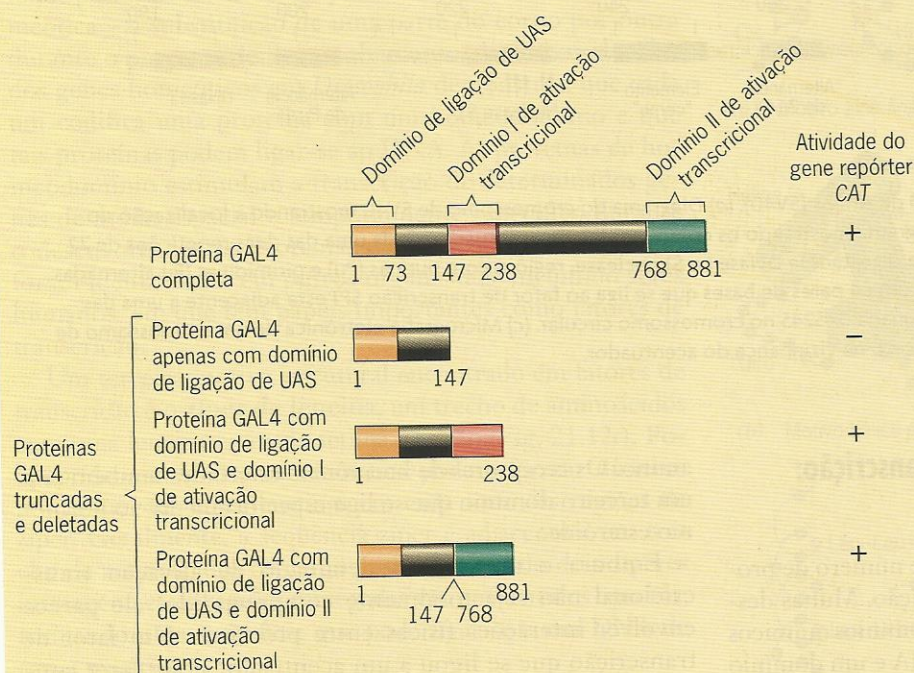
## ENFOQUE TÉCNICO: GAL4, um Fator de Transcrição de Leveduras (continuação)

nômeno foi primeiro demonstrado por Hitoshi Kakidani e Mark Ptashne em experimentos com células de hamster chinês que foram transfectadas com dois plasmídeos, um levando uma versão do gene *GAL4* e o outro levando um gene "repórter" situado posterior às UAS.<sup>1</sup> O gene repórter codificou uma enzima (cloranfenicol-acetiltransferase, CAT) para a qual há disponibilidade de uma análise bioquímica simples. Kakidani e Ptashne determinaram se a proteína *GAL4* expressa por um dos plasmídeos podia estimular a expressão do gene repórter *CAT* no outro plasmídeo ligando-se a UAS próximas. Seus resultados (Fig. 2) mostraram que plasmídeos portadores de um gene *GAL4* intacto, ou um gene *GAL4* que codifica o domínio de ligação de UAS e pelo menos um dos domínios de ativação transcricional da proteína, eram capazes de estimular a expressão do gene repórter

*CAT*. Entretanto, um plasmídeo com um gene *GAL4* que codifica apenas o domínio de ligação de UAS não podia estimular a expressão do gene *CAT*. Assim, os domínios de ligação de UAS e de ativação transcricional da proteína *GAL4* são necessários para a função biológica.

A descoberta de que a proteína *GAL4* pode regular a expressão de um gene perto de uma UAS tornou-se a base de um procedimento para estudar os efeitos da expressão gênica eucariótica *in vivo*. Neste procedimento, uma UAS é inserida antes do sítio de início de transcrição de um gene, e o gene de fusão resultante é introduzido nas células de um organismo, seja transitoriamente por transfecção seja permanentemente por transformação. Um gene *GAL4* funcional é então introduzido nas células para estimular a expressão do gene de fusão. Este procedimento fornece um modo para induzir a expressão do gene de fusão e portanto avaliar o significado da função deste gene em uma determinada população de células.

<sup>1</sup>Kakidani, H., e M. Ptashne. 1988. *GAL4* activates gene expression in mammalian cells. *Cell* 52: 161-167.

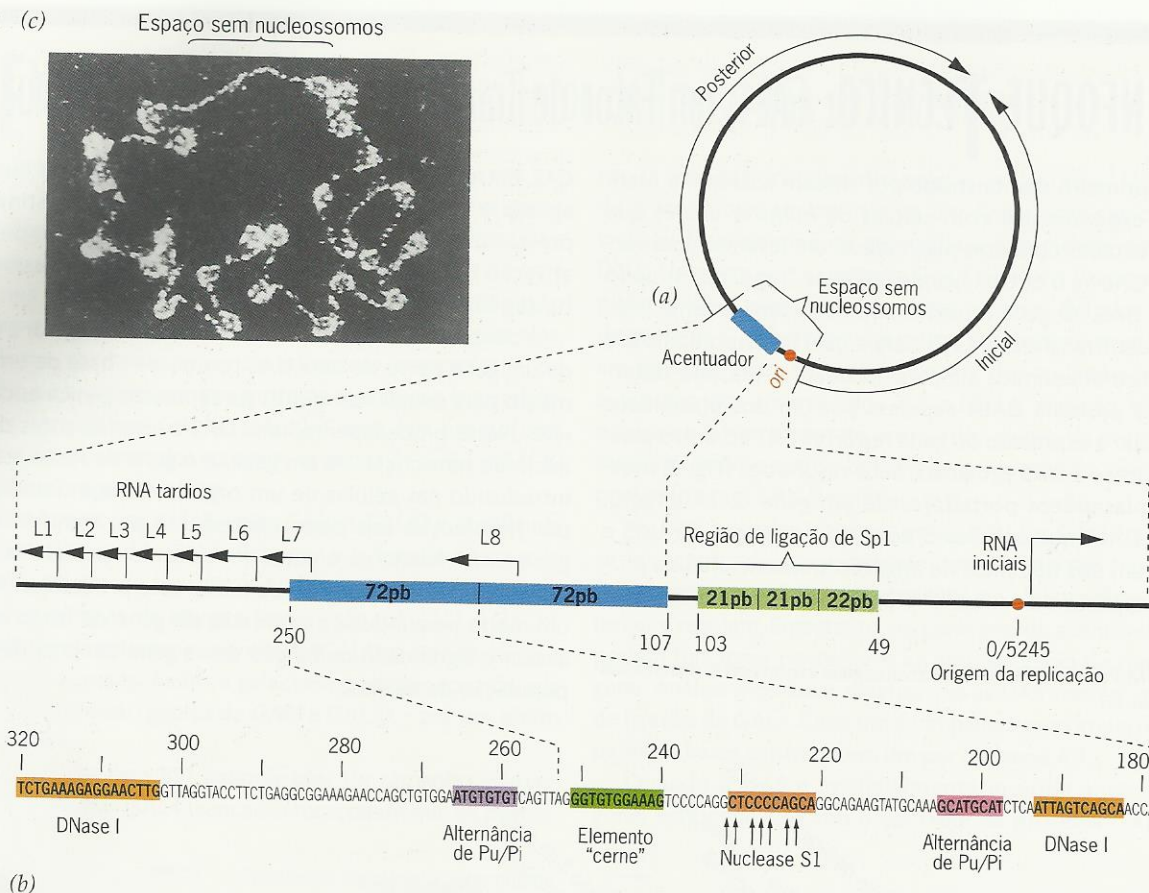


**Fig. 2** ■ Resultados dos experimentos de Kakidani e Ptashne na regulação de *GAL4* de um gene repórter *CAT* fusionado a uma UAS. A proteína *GAL4* completa e duas proteínas deletadas são capazes de estimular a expressão do gene repórter. Todas estas três proteínas *GAL4* contêm o domínio de ligação UAS e pelo menos um dos domínios de ativação transcricional. Uma proteína *GAL4* truncada contendo apenas o domínio de ligação de UAS não estimula a expressão do gene repórter.

Como os acentuadores influenciam a transcrição de genes? Embora não seja conhecida uma resposta detalhada, parece que as proteínas que se ligam a acentuadores influenciam a atividade das proteínas que se ligam a promotores, incluindo os fatores basais de transcrição e a RNA-polimerase. Este achado sugere que os dois grupos de proteínas estão em contato físico, talvez porque os elementos acentuador e promotor ficam juntos por algum tipo de dobramento do DNA. Pesquisas

recentes indicam que um complexo multimérico consistindo em pelo menos 20 proteínas diferentes medeia as interações entre fatores basais de transcrição e as proteínas que se ligam a acentuadores. Este complexo mediador pode dobrar o DNA de tal modo que as proteínas ligadas aos acentuadores fiquem justapostas àquelas ligadas ao promotor. Deste modo, então, proteínas ligadas aos acentuadores podem exercer controle da transcrição, que é iniciado no promotor.





**Fig. 21.11** ■ Estrutura do acentuador do vírus 40 de símios (SV40). (a) Diagrama do cromossomo de SV40 mostrando a localização do acentuador e a origem da replicação (*ori*). (b) Diagrama mostrando os componentes do acentuador. Cada uma das duas repetições de 72 pares de bases contém regiões que são sensíveis à digestão de DNase I e S1 nuclease, regiões com purinas (Pu) e pirimidinas (Pi) alternadas e o elemento acentuador central. Um elemento com 64 pares de bases que se liga ao fator de transcrição SP1 está adjacente a uma das repetições. A origem da replicação está na coordenada 0/5245 no cromossomo circular. (c) Micrografia eletrônica de um cromossomo de SV40 mostrando nucleossomos em toda parte exceto na vizinhança do acentuador.

## Proteínas Envolvidas no Controle da Transcrição: Fatores de Transcrição

Pesquisas recentes identificaram um grande número de proteínas eucarióticas que estimulam a transcrição. Muitas dessas proteínas parecem ter pelo menos dois domínios químicos importantes: um domínio de ligação ao DNA e um domínio de ativação transcricional. Tais domínios podem ocupar partes separadas da molécula ou podem ser superpostos. No fator de transcrição GAL4 de leveduras (veja Enfoque Técnico: GAL4, um Fator de Transcrição de Leveduras), por exemplo, o domínio de ligação ao DNA está situado perto da ponta amino do polipeptídeo. Dois domínios de ativação transcricional estão presentes neste polipeptídeo, um mais ou menos no meio e o outro perto do terminal carboxila. Nas proteínas receptoras de hormônios esteróides, que são fatores de transcrição em animais, o domínio de ligação ao DNA está situado centralmente e parece estar superposto ao domínio de ativação transcricional que vai até o terminal

amino. Os receptores de hormônios esteróides também têm um terceiro domínio que se liga especificamente ao hormônio esteróide.

Embora o mecanismo detalhado de ativação transcricional não seja totalmente compreendido, ele parece envolver interações físicas entre proteínas. Um fator de transcrição que se ligou a um acentuador pode fazer contato com uma ou mais proteínas em outros acentuadores ou pode interagir diretamente com proteínas que estão ligadas à região do promotor. Por meio destes contatos e interações, o domínio de ativação transcricional do fator pode então induzir mudanças conformacionais nas proteínas montadas, abrindo o caminho para a RNA-polimerase iniciar a transcrição.

Muitos fatores de transcrição eucarióticos têm motivos estruturais característicos que resultam de associações entre aminoácidos dentro de suas cadeias polipeptídicas. Um destes motivos é o **zinc finger**, uma curta alça polipeptídica que se forma quando duas cisteínas em uma parte do poli-



peptídeo e duas histidinas em outra parte próxima se juntam a um íon de zinco. O segmento peptídico entre os dois pares de aminoácidos projeta-se do corpo principal da proteína como um tipo de dedo (Fig. 21.12a). Análise mutacional demonstrou que estes dedos têm papéis importantes na ligação do DNA.

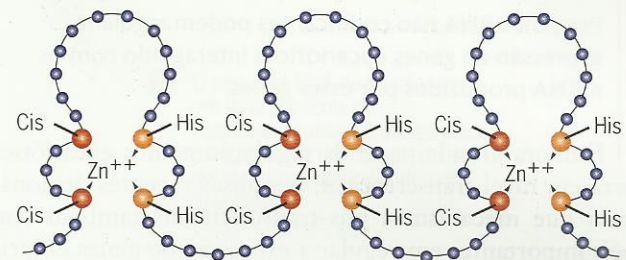
Um segundo motivo em muitos fatores de transcrição é o **hélice-giro-hélice**, um trecho de três hélices pequenas de aminoácidos separadas umas das outras por giros (Fig. 21.12b). Análises genéticas e bioquímicas mostraram que o segmento helicoidal mais próximo do terminal carboxila é necessário para ligação do DNA. As outras hélices parecem estar envolvidas na formação de dímeros de proteína. Em muitos fatores de transcrição, o motivo hélice-giro-hélice coincide com uma região altamente conservada de aproximadamente 60 aminoácidos denominada **homeodomínio**, assim chamada porque ocorre em proteínas codificadas pelos genes homeóticos de *Drosophila*. Análise clássica demonstrou que mutações nestes genes alteram os destinos desenvolvimentais de grupos de células (Cap. 22). Assim, por exemplo, mutações no gene *Antennapedia* podem fazer com que antenas se desenvolvam como pernas. Este fenótipo bizarro é um exemplo de uma transformação homeótica – a substituição de uma parte do corpo por outra durante o processo de desenvolvimento. Análise molecular dos genes homeóticos em *Drosophila* demonstrou que cada um codifica uma proteína com um homeodomínio e que tais proteínas podem ligar-se ao DNA. As proteínas de homeodomínio estimulam a transcrição de determinados genes de modo espacial e temporalmente específico durante o desenvolvimento. Proteínas de homeodomínio também foram identificadas em outros organismos, incluindo seres humanos, em que têm papéis importantes como fatores de transcrição.

Um terceiro motivo estrutural encontrado em fatores de transcrição é o **zíper de leucina**, um trecho de aminoácidos com uma leucina em cada sétima posição (Fig. 21.12c). Polipeptídeos com esta característica podem formar dímeros por interações entre leucinas em cada uma das regiões de zíper. Geralmente, a sequência zíper é adjacente a um trecho de aminoácidos de carga positiva. Quando dois zíperes interagem, essas regiões carregadas vão em sentidos opostos, formando uma superfície que pode ligar-se a um DNA de carga negativa.

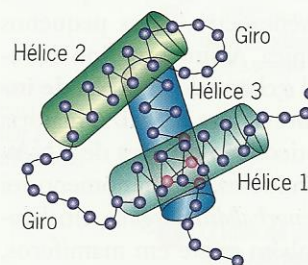
Um quarto motivo estrutural encontrado em alguns fatores de transcrição é a **hélice-alça-hélice**, um trecho de duas regiões helicoidais de aminoácidos separados por uma alça não helicoidal (Fig. 21.12d). As regiões helicoidais permitem a dimerização entre dois polipeptídeos. Às vezes, o motivo hélice-alça-hélice é adjacente a um trecho de aminoácidos básicos (de carga positiva), de modo que, quando ocorre dimerização, esses aminoácidos podem ligar-se ao DNA de carga negativa. Proteínas com esta característica são chamadas *HLH básicas*, ou *bHLH*.

Os fatores de transcrição com motivos de dimerização tais como o zíper de leucina ou hélice-alça-hélice, em princípio,

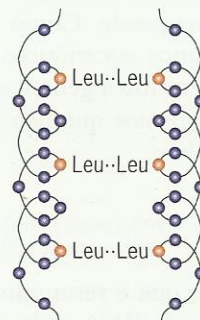
combinam-se com polipeptídeos como eles mesmos para formar homodímeros, ou podem combinar-se com polipeptídeos diferentes para formar heterodímeros. Esta segunda possibilidade sugere um modo pelo qual podem ser atingidos padrões complexos de expressão gênica. A transcrição de um gene em um tecido particular pode depender da ativação por um heterodímero, que só pode formar-se se seus polipeptídeos constituintes forem produzidos neste tecido. Além disso, estes dois polipeptídeos teriam de estar presentes nas quantidades corretas para favorecer a formação do heterodímero em relação ao homodímero correspondente. Modulações sutis na expressão gênica podem portanto ser obtidas mudando-se as concentrações dos dois componentes de um heterodímero.



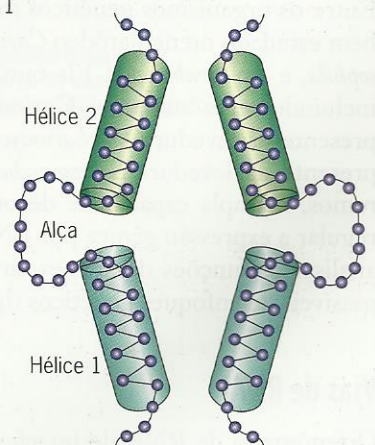
(a) Motivo zinc finger.



(b) Motivo hélice-giro-hélice.



(c) Motivo zíper de leucina.



(d) Motivo hélice-alça-hélice.

**Fig. 21.12** ■ Motivos estruturais dentro de tipos diferentes de fatores de transcrição. (a) Motivos zinc-finger no fator SP1 de transcrição de mamíferos. (b) Motivo hélice-giro-hélice em um fator de transcrição homeodomínio. (c) Motivo zíper de leucina que permite que dois polipeptídeos se dimerizem e então se liguem ao DNA. (d) Um motivo hélice-alça-hélice que permite dimerização de dois polipeptídeos e então a ligação ao DNA.



### PONTOS IMPORTANTES

- Os acentuadores atuam de modo independente de orientação a distâncias consideráveis para regular a transcrição de um promotor do gene.
- Os fatores de transcrição reconhecem seqüências específicas de DNA dentro dos acentuadores e se ligam a elas.
- Os fatores de transcrição possuem motivos estruturais característicos tais como *zinc finger*, hélice-giro-hélice, zíper de leucina e hélice-alça-hélice.

## REGULAÇÃO PÓS-TRANSCRICIONAL DA EXPRESSÃO GÊNICA POR INTERFERÊNCIA DE RNA

Pequenos RNA não codificantes podem regular a expressão de genes eucarióticos interagindo com os mRNA produzidos por estes genes.

Embora grande parte da regulação gênica eucariótica ocorra em nível transcricional, pesquisas recentes demonstraram que mecanismos pós-transcricionais também têm papéis importantes em regular a expressão de genes eucarióticos. Alguns desses mecanismos envolvem pequenos RNA não codificantes. Por pareamento de bases com seqüências alvos em moléculas de RNA mensageiro, esses pequenos RNA interferem na expressão gênica. Assim, este tipo de regulação gênica pós-transcricional é chamado de **RNA de interferência**, geralmente com a abreviatura **RNAi**. A maioria dos tipos de organismos eucarióticos são capazes de RNAi. Entre os organismos genéticos modelos, este fenômeno foi bem estudado no nematódeo *Caenorhabditis elegans*, em *Drosophila*, e em *Arabidopsis*. Ele também existe em mamíferos, incluindo seres humanos. Curiosamente, o RNAi não está presente na levedura *Saccharomyces cerevisiae*, embora esteja presente na levedura *Schizosaccharomyces pombe*. Como veremos, a ampla capacidade de organismos eucarióticos de regular a expressão gênica por RNAi permitiu a geneticistas analisar as funções dos genes em organismos que não são passíveis de enfoques genéticos tipo padrão.

### Vias de RNAi

O fenômeno de RNA de interferência, que é resumido na Fig. 21.13, envolve pequenas moléculas de RNA chamadas de pequenos RNA de interferência (siRNA) ou microRNA (miRNA). Estas moléculas, com 21 a 28 pares de bases de tamanho, são produzidas a partir de moléculas bifilamentares de RNA maiores pela ação enzimática de proteínas que são endonucleases específicas de RNA bifilamentar. Como tais endonucleases cortam grandes RNA em pedaços pequenos, elas são chamadas de enzimas **Dicer**. O nematódeo *Caenorhabditis elegans* produz um único tipo de enzima Dicer,

*Drosophila* produz duas enzimas Dicer diferentes e *Arabidopsis* produz pelo menos três. Em *C. elegans* e *Drosophila*, essas enzimas atuam no citoplasma; em *Arabidopsis*, elas provavelmente atuam no núcleo. Os siRNA e miRNA produzidos por atividade Dicer fazem pares de bases ao longo de seu tamanho exceto em suas pontas 3', onde dois nucleotídeos não estão pareados.

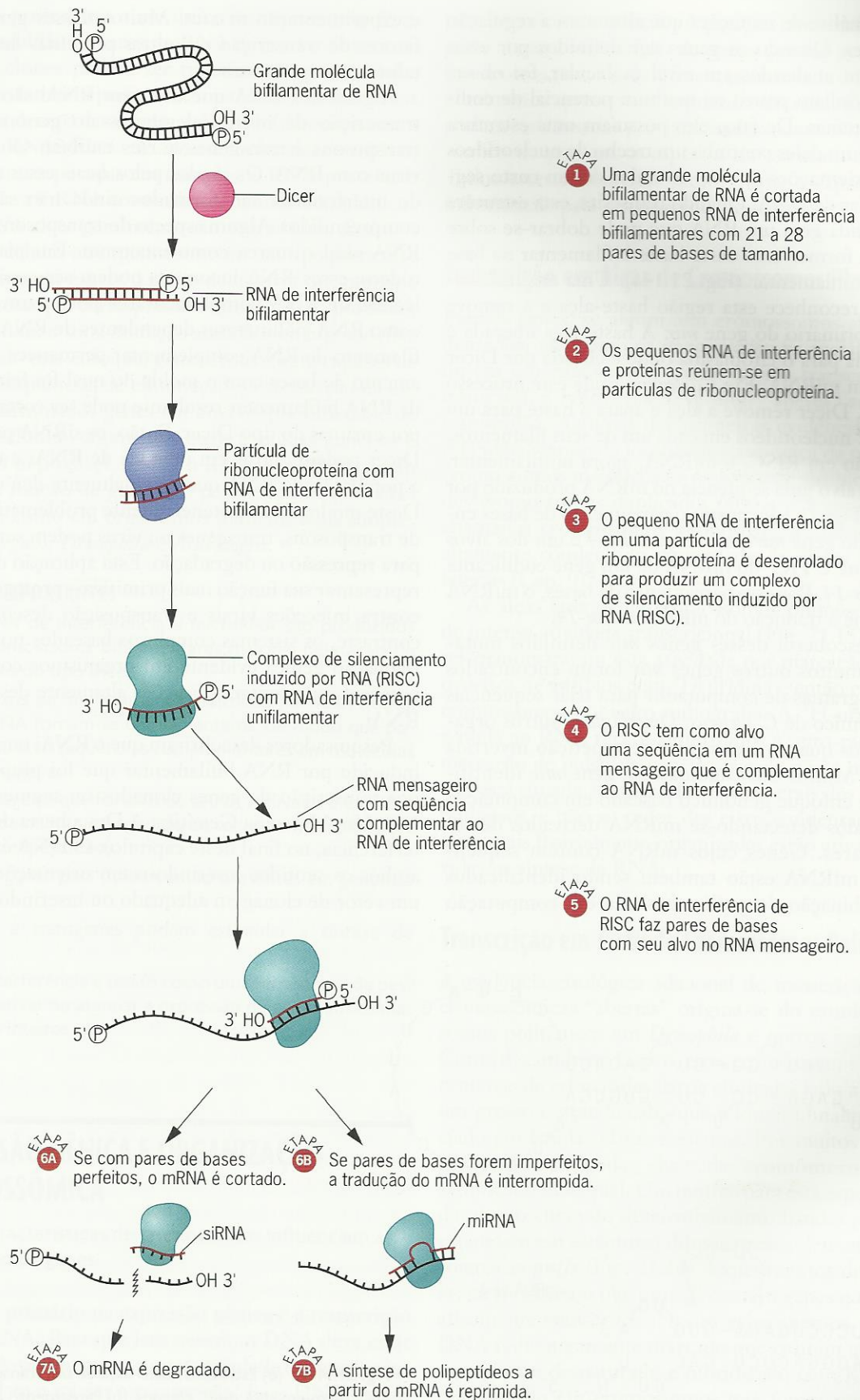
No citoplasma, os siRNA e miRNA tornam-se incorporados a partículas de ribonucleoproteínas. O siRNA bifilamentar ou miRNA nestas partículas é deslicoidizado, e um de seus filamentos é preferencialmente eliminado. O filamento único de RNA sobrevivente é então capaz de interagir com moléculas específicas de RNA mensageiro. Essa interação é mediada por pareamento de bases entre o RNA unifilamentar no complexo RNA-proteína e uma seqüência complementar na molécula de RNA mensageiro. Como esta interação impede a expressão do gene que produziu o mRNA, a partícula RNA-proteína é chamada de **Complexo de Silenciamento Induzido por RNA (RISC)**.

Os RISC de organismos diferentes variam de tamanho e composição. Entretanto, todos contêm pelo menos uma molécula da chamada família Argonauta de proteínas. A função destas proteínas não é totalmente compreendida. Sempre que o pareamento de bases entre o RNA dentro de RISC e a seqüência alvo no mRNA é perfeito ou quase perfeito, o RISC corta o mRNA alvo no meio da região com pareamento de bases. O mRNA cortado é então degradado. A endonuclease responsável por cortar o mRNA atualmente é desconhecida. Entretanto, ela pode ser uma proteína Argonauta. Após a clivagem, o RISC pode associar-se a outra molécula de mRNA e induzir sua clivagem. Como RISC pode ser usado repetidamente sem perder sua capacidade de atingir e cortar o mRNA, ele se comporta como um catalisador. RNA associados ao RISC que resultam na clivagem do mRNA geralmente são chamados de **pequenos RNA de interferência**. Sempre que o RNA dentro do RISC fizer par imperfeitamente com sua seqüência alvo, o mRNA geralmente não é clivado; em vez disso, a tradução do mRNA é inibida. RNA associados ao RISC que têm este efeito geralmente são chamados de **microRNA**. Em animais, as seqüências alvos de RISC são encontradas nas regiões 3' não traduzidas de moléculas de mRNA, e geralmente tais seqüências estão presentes várias vezes dentro da região 3' não traduzida (UTR). Em plantas, as seqüências alvos de RISC geralmente estão situadas dentro da região codificante do mRNA ou dentro das UTR 5' dos mRNA.

### Fontes de Pequenos RNA de Interferência e microRNA

Algumas das moléculas do pequeno RNA que induzem o RNAi são derivadas de transcritos dos genes de microRNA. Estes genes, em geral indicados pelo símbolo *mir*, são encontrados nos genomas de muitos tipos de eucariotes; cerca de 100 genes *mir* estão presentes nos genomas de *C. elegans* e *Drosophila* e cerca de 250 estão presentes nos genomas de vertebrados. Inicialmente, alguns destes genes foram iden-





**Fig. 21.13** ■ Resumo dos eventos envolvidos nas vias de RNA de interferência.



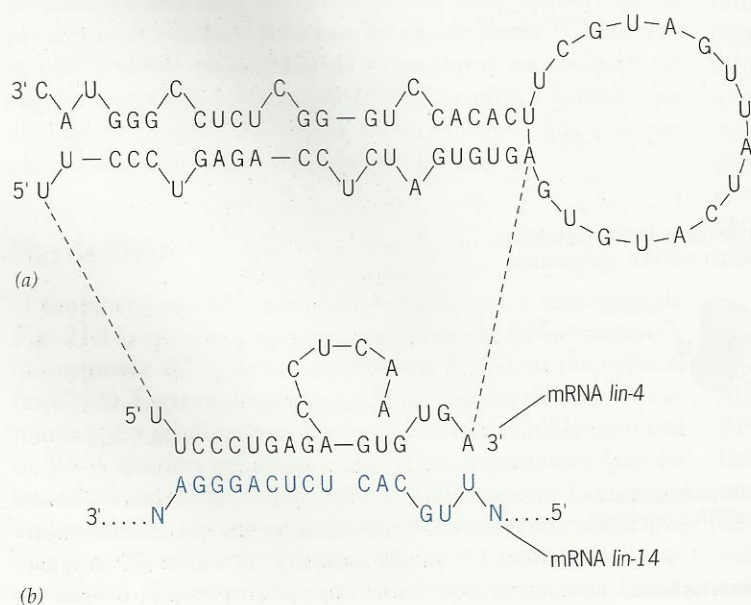
tificados por análise de mutações que alteraram a regulação de outros genes. Quando os genes *mir* definidos por essas mutações foram analisados em nível molecular, foi observado que eles tinham pouco ou nenhum potencial de codificação de proteínas. De fato, eles possuíam uma estrutura peculiar. Cada um deles continha um trecho de nucleotídeos repetido em orientações opostas ao redor de um curto segmento intercalar de DNA. Quando transcrita, esta estrutura invertida repetida gera um RNA que pode dobrar-se sobre si mesmo para formar uma curta haste bifilamentar na base de uma alça unifilamentar (Fig. 21.14a). Uma enzima chamada Droscha reconhece esta região haste-alça e a remove do transcrito primário do gene *mir*. A haste-alça liberada é então exportada para o citoplasma onde é clivada por Dicer para formar um miRNA. Em *C. elegans*, onde este processo foi descoberto, Dicer remove a alça e apara a haste para um tamanho de 22 nucleotídeos em cada um de seus filamentos. Após maturação em RISC, o miRNA, agora unifilamentar, pode ter como alvo uma sequência no mRNA produzido por outro gene. A Fig. 21.14b mostra o pareamento de bases entre o miRNA do gene *mir* de *C. elegans lin-4* e um dos alvos deste miRNA na UTR3' do mRNA de um gene codificante de proteína, *lin-14*. Por este pareamento de bases, o miRNA de *lin-4* reprime a tradução do mRNA de *lin-14*.

Desde a descoberta desses genes *mir* definidos mutacionalmente, muitos outros genes *mir* foram encontrados usando-se programas de computador para triar sequências de DNA genômico de *C. elegans*, *Drosophila* e outros organismos modelos quanto à estrutura de repetição invertida característica. Muitos dos candidatos a gene *mir* identificados por este enfoque genômico baseado em computação foram verificados detectando-se miRNA derivados desses extratos celulares. Genes cujos mRNA contêm sequências alvos de miRNA estão também sendo identificados por uma combinação de análise baseada em computação

e experimentação *in vivo*. Muitos desses genes codificam fatores de transcrição ou outras proteínas desenvolvimentalmente importantes.

Alguns dos RNA que induzem RNAi são derivados da transcrição de outros elementos no genoma, tais como transposons e transgenes, e eles também são derivados de vírus com RNA. Os modos pelos quais esses tipos de RNA de interferência são formados ainda não são totalmente compreendidos. Algum aspecto do transposon, transgene ou RNA viral o marca como incomum. Em plantas e nematódeos, esses RNA incomuns podem ser copiados em moléculas de RNA complementares por enzimas conhecidas como RNA-polimerases dependentes de RNA (RdRP). Se o filamento de RNA complementar permanece fazendo pareamento de bases com o molde do qual foi feito, a molécula de RNA bifilamentar resultante pode ser cortada em siRNA por enzimas do tipo Dicer. Então, os siRNA produzidos por Dicer podem entrar em uma via de RNAi e ter como alvo a população de RNA que originalmente deu origem a eles. Deste modo, RNA potencialmente problemáticos derivados de transposons, transgenes ou vírus podem ser direcionados para repressão ou degradação. Esta aplicação de RNAi pode representar sua função mais primitiva – proteger organismos contra infecções virais e transposição descontrolada. Em contraste, os sistemas complexos baseados no miRNA para regulação gênica evidente em organismos como *C. elegans* parecem representar aplicações altamente desenvolvidas do RNAi.

Pesquisadores descobriram que o RNAi também pode ser induzido por RNA bifilamentar que foi preparado *in vitro* por transcrição de genes clonados ou segmentos de genes (veja Um Marco na Genética: A Descoberta do RNA de Interferência, no final deste capítulo). O DNA é transcrito em ambos os sentidos inserindo-o em orientações opostas em um vetor de clonagem adequado ou inserindo cópias inver-



**Fig. 21.14** ■ (a) Estrutura haste-alça de um transcrito do gene *lin-4* de microRNA de *C. elegans*. (b) Pareamento de bases entre o microRNA derivado do transcrito de *lin-4* e uma sequência na região 3' não traduzida do RNA mensageiro de *lin-14*.



tidas do DNA posteriores a um único promotor (veja Cap. 17). Moléculas bifilamentares de RNA derivadas dos transcritos de tais clones podem ser transfectadas para células cultivadas. Elas também podem ser injetadas em organismos vivos. Uma vez dentro das células, o RNA bifilamentar entra em uma via de RNAi. Ele é cortado em moléculas de siRNA, que são então incorporadas em complexos RNA-proteína e direcionadas para mRNA contendo seqüências complementares. Os mRNA alvos geralmente são degradados. Assim, o tratamento de células ou organismos com um tipo particular de RNA bifilamentar exerce o efeito de inativar ou atenuar a expressão do gene que corresponde a este RNA. É, portanto, equivalente a induzir uma mutação amórfica ou hipomórfica no gene. Usando este enfoque, geneticistas foram capazes de estudar as conseqüências de eliminação ou atenuação da expressão de determinados genes em uma ampla variedade de organismos, incluindo alguns nos quais a análise genética é difícil, lenta ou impossível. Assim, o RNAi está sendo usado para analisar a função de genes em peixes, roedores e seres humanos, bem como em organismos modelos mais simples tais como *C. elegans*, *Drosophila* e *Arabidopsis*.

#### ■ PONTOS IMPORTANTES

- Pequenos RNA de interferência e microRNA são produzidos a partir de precursores bifilamentares maiores pela ação de endonucleases do tipo Dicer.
- Nos Complexos de Silenciamento Induzidos por RNA (RISC), siRNA e miRNA tornam-se unifilamentares, de modo que podem ter como alvo seqüências complementares em moléculas de RNA mensageiro.
- O RNA mensageiro que foi alvo de siRNA é clivado e o mRNA que foi alvo de miRNA é impedido de servir como molde para a síntese de polipeptídeos.
- Centenas de genes para miRNA estão presentes em genomas eucarióticos.
- Transposons e transgenes podem estimular a síntese de siRNA.
- O RNA de interferência é usado como uma ferramenta de pesquisa para inativar ou atenuar a expressão de genes em células e organismos inteiros.

## EXPRESSION GÊNICA E ORGANIZAÇÃO CROMOSSÔMICA

Muitas características de cromossomos influenciam a transcrição de genes.

O evento primário na expressão gênica é a transcrição do DNA em RNA. Para que isto ocorra, o DNA deve estar acessível à RNA-polimerase e a uma variedade de outras proteínas que ajudam a iniciar a transcrição. Se o DNA estiver fortemente ligado a histonas ou a outros tipos de proteínas “de embalagem”, ele também pode estar muito condensado

para permitir transcrição. A posição cromossômica também pode afetar a transcrição de genes porque genes que foram transpostos para locais diferentes dentro de um cromossomo geralmente apresentam expressão alterada. Essas alterações podem ser causadas por acentuadores no novo local cromossômico ou por outros aspectos da estrutura cromossômica. Todos estes fenômenos indicam que a expressão gênica é influenciada pela organização cromossômica.

### Transcrição em Alças de Cromossomos Plumosos

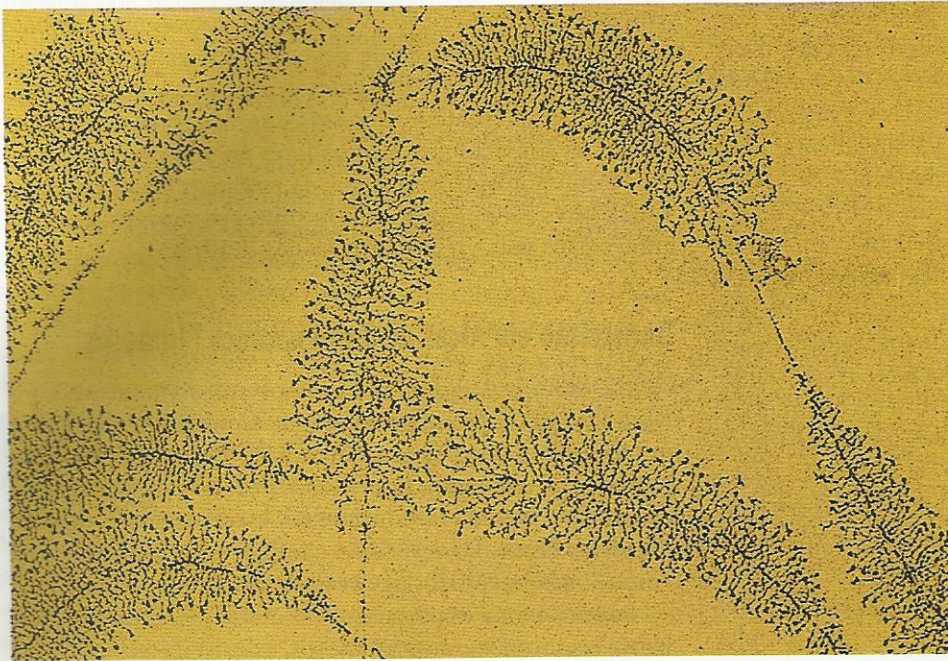
Poucos genes, se algum, em cromossomos altamente condensados, tais como os encontrados na metáfase da mitose ou em espermatozoides maduros, são transcritos. Para ser transcrita, a cromatina deve estar “aberta” ao aparelho de transcrição. Algumas das primeiras evidências que apóiam esta idéia vieram da localização citológica da transcrição nos cromossomos meióticos de ovócitos de anfíbios. Estes grandes cromossomos duplicados (400 a 800 µm de tamanho), chamados cromossomos plumosos, consistem em um eixo altamente condensado circundado de vários pares de alças laterais (Cap. 9).

As alças laterais de cromossomos plumosos são regiões de intensa atividade transcricional (Fig. 21.15). Pesquisadores demonstraram esta atividade por marcação de pulso em ovócitos de anfíbios com <sup>3</sup>H-uridina e então examinando os cromossomos por auto-radiografia. A <sup>3</sup>H-uridina é incorporada ao RNA recém-sintetizado. A auto-radiografia com marcação de pulso de ovócitos revela que a uridina radioativa está situada ao redor das alças laterais dos cromossomos plumosos, e não ao redor dos eixos condensados. Assim, as alças mais frouxamente organizadas estão ativamente envolvidas na síntese de RNA.

### Transcrição em Puffs de Cromossomos Politênicos

A evidência citológica adicional de transcrição em regiões cromossômicas “abertas” origina-se do estudo de cromossomos politênicos em *Drosophila* e outros insetos dípteros. Como discutido no Cap. 6, estes cromossomos consistem em centenas de cromátides irmãs alinhadas lado a lado, criando um grosso e grande cabo que é longitudinalmente diferenciado em bandas claras e escuras. Por muitos anos, acreditava-se que as bandas, chamadas **cromômeros**, tinham um significado funcional. Um motivo para esta especulação é que durante o curso do desenvolvimento bandas particulares se expandem em estruturas difusas menos densamente coradas chamadas **puffs** (Fig. 21.16). Experimentos de hibridização *in situ* mostraram que os **puffs** contêm genes que estão sendo ativamente transcritos. Nestes experimentos, um RNA ou DNA radioativamente marcado que contém uma seqüência específica é desnaturado e hibridizado com RNA ou DNA desnaturado em cromossomos que foram preparados para análise citológica. A sonda marcada de RNA ou DNA liga-se à sua seqüência complementar nos cromossomos, isto é,





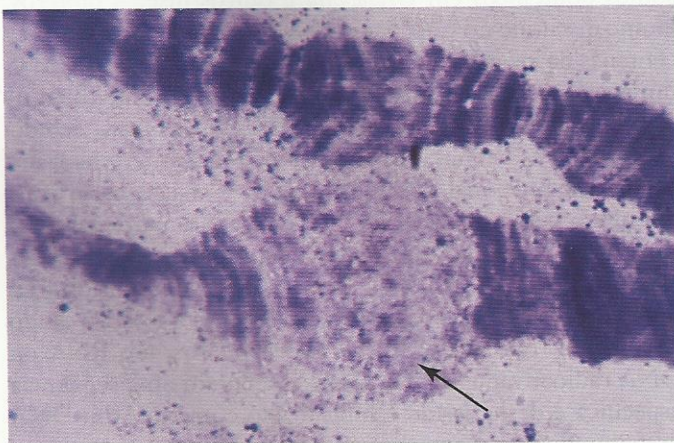
**Fig. 21.15** ■ Fotomicrografia mostrando a transcrição das alças de DNA de cromossomos plumosos de anfíbio.

liga-se *in situ*. A localização da sonda ligada é então revelada por auto-radiografia. A aplicação desta técnica mostra que os *puffs* em cromossomos politênicos contêm seqüências de DNA que estão sendo ativamente transcritas em RNA. Com o tempo, determinados *puffs* regredem e formam-se novos em locais diferentes nos cromossomos. Esta seqüência temporal de formação de *puffs* é controlada pela *ecdisona*, um hormônio esteróide que aparece em pulsos em épocas diferentes durante o desenvolvimento. Uma interpretação plausível é que a *ecdisona* induz transcrição em bandas com *puffs*. *Puffs* também podem ser induzidos por outros fatores como choque provocado por calor. Quando larvas de *Drosophila* são incubadas em temperaturas maiores que 33°C, vários novos *puffs* aparecem nos cromossomos politênicos, e os anteriores regredem. Os novos *puffs* correspondem aos genes do choque térmico. Tais estudos citológicos com *Drosophila*

mostram, portanto, que regiões cromossômicas difusas expandidas são sítios de intensa síntese de RNA.

### Organização Molecular de DNA Transcricionalmente Ativo

Os achados com cromossomos plumosos e politênicos levantam questões sobre a organização molecular do DNA transcricionalmente ativo. Esse DNA é de fato mais “aberto” que o DNA não transcrito? Esta pergunta foi respondida avaliando-se a sensibilidade do DNA na cromatina à ação da desoxirribonuclease I pancreática (DNase I), uma enzima que corta moléculas de DNA e as degrada em seus nucleotídeos constituintes. Em 1976, Mark Groudine e Harold Weintraub demonstraram que o DNA transcricionalmente ativo é mais sensível à DNase I do que o DNA não transcrito. Groudine e Weintraub extraíram cromatina de hemácias de galinhas e a digeriram parcialmente com DNase I. Então eles sondaram o material residual da cromatina quanto a seqüências de dois genes,  $\beta$ -globina, que é ativamente transcrito em hemácias, e ovalbumina, que não é. Eles descobriram que mais de 50% do DNA de  $\beta$ -globina tinha sido digerido pela enzima DNase I, em comparação com apenas 10% do DNA de ovalbumina. Estes resultados sugerem fortemente que o gene ativamente transcrito estava mais “aberto” ao ataque da nuclease. Pesquisas subseqüentes mostraram que a sensibilidade à nuclease de genes transcricionalmente ativos depende de pelo menos duas pequenas proteínas não-histonas, HMG14 e HMG17 (HMG de grupo de alta mobilidade, pois têm alta mobilidade durante a eletroforese em gel). Quando tais proteínas são removidas da cromatina ativa, a sensibilidade à nuclease é perdida; quando são novamente adicionadas, ela é restaurada.



**Fig. 21.16** ■ Um puff (seta) nos cromossomos politênicos de *Drosophila*.



O tratamento da cromatina isolada com uma concentração muito baixa de DNase I faz com que o DNA seja clivado em alguns sítios específicos, apropriadamente chamados de **sítios hipersensíveis à DNase I**. Foi demonstrado que alguns desses sítios estão antecedentes a genes transcricionalmente ativos, em regiões do promotor ou do acentuador. O significado funcional destes sítios hipersensíveis ainda não está claro, mas algumas evidências sugerem que podem marcar regiões nas quais o DNA está localmente deselcoicidizado, talvez porque a transcrição começou.

No caso dos genes de  $\beta$ -globina humana, vários sítios hipersensíveis à DNase I estão situados em uma **região controladora de locus (LCR)** com 15 kb de tamanho antecedente aos próprios genes. A LCR contém sítios de ligação para fatores de transcrição que pré-ativam os genes de  $\beta$ -globina para transcrição. A pré-ativação é detectada por aumento na sensibilidade do DNA dentro da LCR a digestão com baixas concentrações de DNase I. A transcrição dos genes de  $\beta$ -globina parece requerer essa pré-ativação e é estimulada por fatores de transcrição que se ligam a acentuadores específicos no complexo do gene de  $\beta$ -globina. Entretanto, o tecido e a especificidade temporal da expressão do gene de  $\beta$ -globina dependem de seqüências inseridas na LCR. Estudos com camundongos transgênicos indicam que a LCR não é apenas uma grande coleção de acentuadores que exercem controle sobre vários genes de  $\beta$ -globina. A LCR deve estar situada antecedente aos genes de  $\beta$ -globina e em sua orientação natural para controlar apropriadamente a expressão gênica. Isto é, ela funciona de modo dependente da orientação. Os acentuadores tipicamente funcionam de modo independente de orientação e em posições diferentes em relação a um promotor do gene. A LCR tem outra característica que a distingue dos acentuadores simples: ela pode controlar a expressão do gene de  $\beta$ -globina quando todo o grupo de genes (LCR mais os genes de  $\beta$ -globina) está inserido em uma posição cromossômica diferente. Os acentuadores, em contraste, em geral falham em funcionar quando eles e seus genes associados são transpostos para um local cromossômico diferente. Assim, a LCR parece isolar os genes de  $\beta$ -globina da influência da cromatina ao redor deles.

## Remodelagem da Cromatina

Experimentos que avaliam a sensibilidade do DNA a digestão com DNase I estabeleceram que o DNA transcrito é mais acessível ao ataque da nuclease do que o DNA não transcrito. O DNA transcrito está embalado em nucleossomos? Se está, que mudanças estruturais ocorrem nos nucleossomos durante a transcrição? Os nucleossomos estão "abertos" e "fechados" à medida que a RNA-polimerase passa pelo DNA molde? Os esforços para responder a estas perguntas têm envolvido uma combinação de enfoques genéticos e bioquímicos segundo os quais o DNA transcrito está de fato compactado em nucleossomos. Entretanto, no DNA transcrito, os nucleossomos estão alterados por complexos multiprotéticos que facilitam a ação da RNA-polimerase. Esta

alteração de nucleossomos na preparação para transcrição é chamada de **remodelagem da cromatina**.

Foram identificados dois tipos gerais de complexos de remodelagem de cromatina. Um tipo é composto de enzimas que transferem grupos acetil para o aminoácido lisina em posições específicas nas histonas dos nucleossomos. Como uma classe, estas enzimas são chamadas de histona-acetiltransferases (HAT). Vários estudos mostraram que a acetilação de histonas está correlacionada com aumento de expressão gênica, talvez porque a adição de grupos acetil afrouxa a associação entre o DNA e os octâmeros de histona nos nucleossomos. As cinases, enzimas que transferem grupos fosfato para moléculas, também podem ter um papel junto com esses complexos de remodelagem de cromatina. Sabe-se, por exemplo, que a acetilação de lisina-14 em histona H4 em geral é precedida de fosforilação de serina-10 nesta molécula. Juntas, estas duas modificações da histona H4 parecem "abrir" a cromatina para aumento de atividade transcricional.

Um outro tipo de complexo de remodelagem de cromatina perturba a estrutura do nucleossomo na vizinhança do promotor de um gene. O mais intensamente estudado destes complexos é o SWI/SNF encontrado na levedura do pão. Este complexo é designado pelos dois tipos de mutações (mudança inibida e não fermentador de sacarose) que levaram à descoberta de suas proteínas constituintes. Complexos correlatos foram encontrados nas células de outros organismos, incluindo seres humanos. O complexo SWI/SNF consiste em pelo menos oito proteínas. Ele regula a transcrição deslocando octâmeros de histona ao longo do DNA associado em nucleossomos. Ele também pode transferir esses octâmeros para outros locais em uma molécula de DNA. A mudança de nucleossomos catalisada pelo complexo SWI/SNF aparentemente dá aos fatores de transcrição acesso ao DNA. Tais fatores então estimulam a expressão do gene.

Discutimos a remodelação da cromatina do ponto de vista da ativação gênica. Entretanto, a cromatina ativa também pode ser remodelada em cromatina inativa. Essa remodelagem reversa parece envolver duas modificações bioquímicas das histonas nos nucleossomos: desacetilação, catalisada pelas histona-desacetilases (HDAC), e metilação, catalisada pelas histona-metiltransferases (HMT). Alguns dos nucleotídeos no DNA também podem ser metilados por um grupo de enzimas chamadas DNA-metiltransferases (DNMT). A cromatina que foi submetida a estas modificações tende a ser transcricionalmente silenciosa. A metilação do DNA é explorada com mais detalhes em outra seção desta obra.

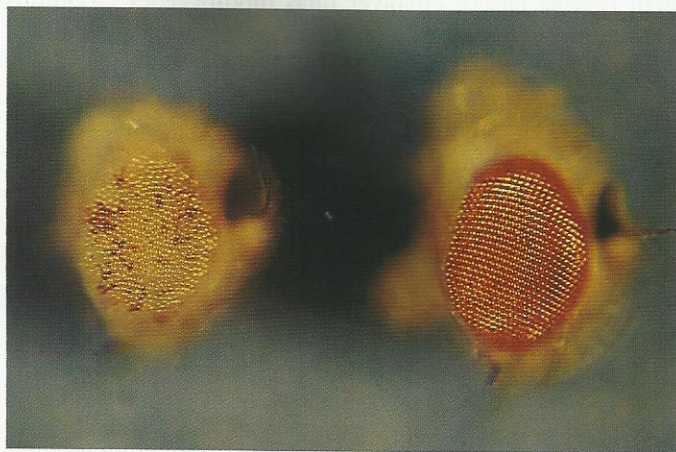
## Eucromatina e Heterocromatina

Varição na densidade da cromatina dentro dos núcleos das células resulta em uma coloração diferencial de trechos dos cromossomos. O material densamente corado é chamado de **heterocromatina**, e sua contraparte de coloração clara é chamada de **eucromatina**. Qual o significado funcional destes tipos diferentes de cromatina?



Uma combinação de análise genética e molecular mostrou que a maioria dos genes eucarióticos estão situados na eucromatina. Além disso, quando genes eucromáticos são artificialmente transpostos para um ambiente heterocromático, eles tendem a funcionar anormalmente, e, em alguns casos, não funcionar. Essa capacidade prejudicada de funcionar pode criar uma mistura de características normais e mutantes no mesmo indivíduo, uma condição chamada de **variação de efeito de posição**. Este termo é usado porque a variabilidade no fenótipo é causada por mudança de posição do gene eucromático, especificamente relocando-o para a heterocromatina. Muitos exemplos de variação de efeito de posição foram descobertos em *Drosophila*, geralmente em associação a inversões ou translocações que movem um gene eucromático para a heterocromatina. O alelo *white mottled*,  $w^{m4}$ , é um bom exemplo. Neste caso, um alelo tipo selvagem do gene *white* foi relocado por uma inversão, com uma quebra perto do locus *white* e a outra na heterocromatina basal do cromossomo X. Este rearranjo interfere na expressão normal do gene *white* e causa um fenótipo olho salpicado. Aparentemente, o gene eucromático *white* não pode funcionar bem em um ambiente heterocromático. Este e outros exemplos levaram ao conceito de que a heterocromatina reprime o funcionamento gênico, talvez porque está condensada em uma forma que não é acessível à maquinaria transcricional. Pesquisas continuadas tentam identificar as proteínas que podem estar envolvidas neste processo de condensação.

Uma estratégia de pesquisa é identificar genes que, quando mutantes, suprimem ou acentuam a variação do efeito de posição. Triagens de tais mutações usaram o alelo variegado  $w^{m4}$  do locus *white* em *Drosophila* como uma análise conveniente. Quando  $w^{m4}$  é combinado com um supressor de variação do efeito de posição, isto é, uma mutação *su(var)*, a cor de olho da mosca será tipo selvagem, ou quase; quando  $w^{m4}$  é combinado com um acentuador de variação do efeito de posição, isto é, com uma mutação *e(var)*, a cor do olho será mais salpicada do que é com  $w^{m4}$  isolado. A Fig. 21.17



**Fig. 21.17** ■ Fotos mostrando o fenótipo do olho de uma mosca com o alelo  $w^{m4}$  (à esquerda) e uma mosca com este alelo junto com um supressor de variação dominante, *Su(var)205*<sup>4</sup> (à direita).

mostra o efeito de um dominante supressor da variação do efeito de posição, *Su(var)205*<sup>4</sup>, no fenótipo  $w^{m4}$ . Análise molecular demonstrou que esta mutação supressora impede a produção de uma proteína cromossômica não histônica que se associa predominantemente com a heterocromatina. Moscas heterozigotas para a mutação *Su(var)205*<sup>4</sup> suprimem o fenótipo  $w^{m4}$ . Entretanto, moscas que são homozigotas para ela morrem cedo no desenvolvimento. Assim, a proteína codificada pelo gene tipo selvagem *Su(var)205*, chamada *proteína de heterocromatina 1*, ou simplesmente *HP1*, é essencial para a vida. Talvez uma mutação em uma cópia do gene *Su(var)205* reduza a quantidade de HP1 o suficiente para relaxar a heterocromatina mas não o suficiente para a desenrolar. Neste estado relaxado, um gene eucromático tal como  $w^{m4}$  que ficou sob a influência repressiva da heterocromatina é expresso em níveis normais. Consequentemente, o fenótipo que ele controla pára de variegar.

Aproximadamente 30% do DNA de *Drosophila* é heterocromático. A maioria desse DNA está situada ao redor dos centrômeros dos cromossomos. Entretanto, parte dele está na eucromatina e parte situada nas pontas, ou telômeros, dos cromossomos. Assim, em *Drosophila*, estão presentes três tipos de heterocromatina: heterocromatina *êntrica* (ao redor dos centrômeros), heterocromatina *intercalar* (dispersa na eucromatina) e heterocromatina *telomérica* (nas pontas dos cromossomos).

A maioria das regiões heterocromáticas no genoma de *Drosophila* são ricas em seqüências de DNA repetitivo. Tanto seqüências altamente repetidas quanto moderadamente repetidas podem ser encontradas nestas regiões, e as seqüências moderadamente repetidas compreendem muitos tipos de elementos de transposição, especialmente retroposons inativos (Cap. 18). As regiões heterocromáticas do genoma de *Drosophila* também contêm alguns genes expressos. Várias centenas de genes para RNA ribossômico estão situadas na heterocromatina *êntrica* dos cromossomos X e Y, e sabe-se que cerca de 20 genes codificantes de proteínas estão em outras regiões heterocromáticas do genoma de *Drosophila*. Assim, a heterocromatina nem sempre está associada a inatividade genética completa. Curiosamente, quando um gene heterocromático é levado para um ambiente eucromático por um rearranjo cromossômico, sua expressão tende a variegar. Além disso, essa variação tende a ser acentuada pelas mutações *su(var)* que suprimem a variação de um gene eucromático que foi colocado em um ambiente heterocromático. O fato de que genes heterocromáticos exibem efeitos de posição opostos aos genes eucromáticos indica que eles estão especialmente adaptados a funcionar em um ambiente heterocromático. Entretanto, a natureza dessa adaptação atualmente é desconhecida.

## Silenciamento Gênico

Embora a heterocromatina tenha a capacidade de reprimir os genes eucromáticos que foram colocados próximo a ela por um rearranjo cromossômico, esta característica da cromatina



não é a única que impede a expressão gênica em eucariontes. Estudos com muitos organismos diferentes indicam que outros aspectos de organização da cromatina, alguns possivelmente relacionados com os componentes protéicos da heterocromatina, controlam a atividade gênica. Nós já vimos que fatores de transcrição, agindo em conjunto com seqüências acentuadoras e promotoras particulares, ligam genes de modo temporal e espacialmente específico. Entretanto, em determinada época de uma célula eucariótica, a maioria dos genes não estão sendo transcritos. Assim, a regra, e não a exceção, é que os genes estão transcricionalmente silenciosos. Que características de organização da cromatina causam este amplo silenciamento gênico?

### Silenciamento pelas Proteínas do Grupo *Polycomb*

Em *Drosophila*, estudos de genes que controlam a formação de determinadas partes do corpo identificaram proteínas que estão envolvidas em silenciamento gênico. Tais proteínas são codificadas por uma classe especial de genes conhecidos como grupo *Polycomb*, abreviado por *PcG*. A denominação desta classe se deve a *Polycomb*, um gene definido por uma mutação dominante que causa a formação de cerdas extras em grupos nas pernas. Esses grupos, conhecidos como pentes sexuais, são normalmente encontrados nas pernas anteriores do macho de *Drosophila*. Entretanto, em machos que levam uma mutação *Polycomb*, eles são encontrados nas pernas média e posterior também. Mutações *Polycomb* também causam outras anomalias fenotípicas, incluindo a transformação de antenas em pernas rudimentares.

*Polycomb* e outros genes *PcG* codificam proteínas que regulam a expressão de fatores de transcrição envolvidos no desenvolvimento de *Drosophila*. Estas proteínas reguladoras parecem formar grandes complexos multiprotéicos que se associam a regiões específicas do genoma – em particular, os loci que codificam tais fatores de transcrição desenvolvimentalmente importantes. Quando e a maneira como fazem isto permanecem um mistério. Entretanto, são conhecidos vários fatos. Primeiro, as proteínas *PcG* associam-se a regiões cromossômicas que contêm seqüências especiais de DNA conhecidas como *elementos de resposta Polycomb* (*PRE*). Individualmente, a maioria das proteínas *PcG* não são capazes de se ligar ao próprio DNA *PRE*. Entretanto, podem ser capazes de se ligar a um complexo, ou podem interagir com outros tipos de proteínas que podem ligar-se ao DNA *PRE*. Segundo, as proteínas *PcG* reprimem os genes situados perto dos *PRE*. Se as proteínas *PcG* se associam ou não a um *PRE* particular e reprimem os genes em sua vizinhança depende de fatores presentes no animal bem cedo em seu desenvolvimento embrionário. Terceiro, uma vez estabelecida, a associação entre as proteínas *PcG* e um *PRE* é fielmente mantida durante o desenvolvimento. Assim, o estado reprimido dos genes perto de um *PRE* é transmitido de uma célula para suas filhas a cada divisão mitótica. Esta forma de silenciamento gênico assegura, portanto, que clones de célu-

las se desenvolvam consistentemente em determinadas partes do corpo. Quando o mecanismo *PcG* de silenciamento gênico está comprometido por mutações em qualquer dos genes *PcG*, algumas partes do corpo desenvolvem-se inapropriadamente, como, por exemplo, no próprio mutante *Polycomb*, em que antenas podem desenvolver-se em pernas e as pernas média e posterior podem desenvolver-se como pernas anteriores.

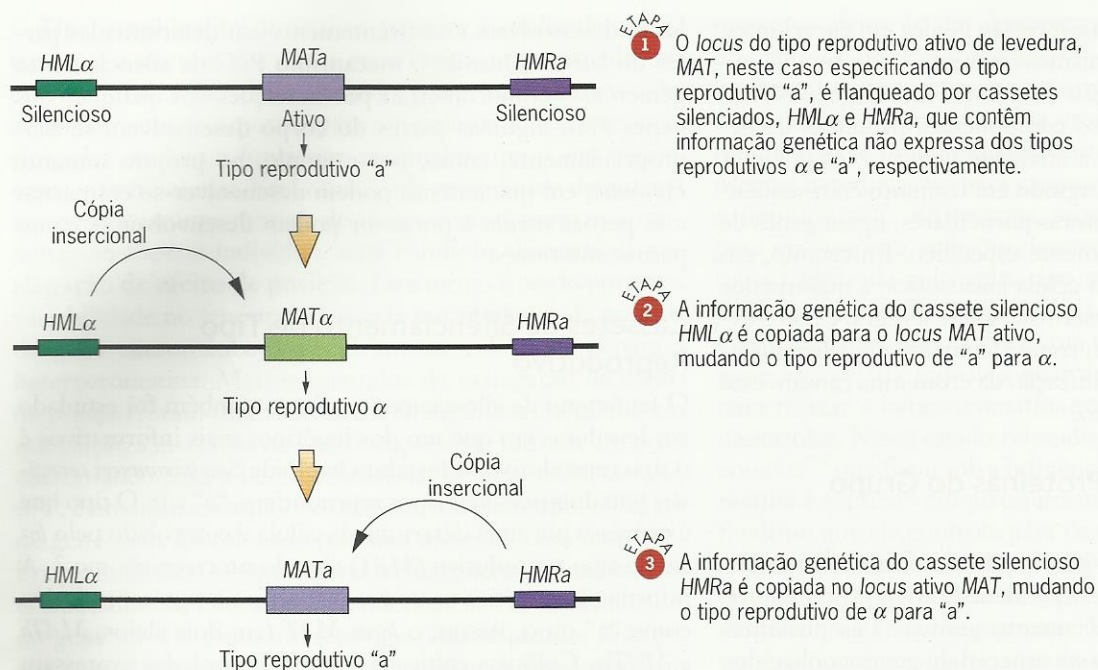
### Cassetes de Silenciamento de Tipo Reprodutivo

O fenômeno de silenciamento gênico também foi estudado em leveduras em que um dos fenótipos mais informativos é o tipo reprodutivo. A levedura haplóide *Saccharomyces cerevisiae* tem dois possíveis tipos reprodutivos, “a” e  $\alpha$ . O tipo que é expresso por uma determinada célula é controlado pelo locus do tipo reprodutivo (*MAT*) situado no cromossomo 3. A informação genética nesse locus especifica o tipo reprodutivo como “a” ou  $\alpha$ . Assim, o locus *MAT* tem dois alelos, *MAT<sub>a</sub>* e *MAT $\alpha$* . Colônias cultivadas de células isoladas expressam todas a mesma informação do tipo reprodutivo. Entretanto, em uma frequência que depende da linhagem particular de leveduras, o tipo reprodutivo pode mudar de “a” para  $\alpha$ , ou de  $\alpha$  para “a”.

A base para essa mudança hoje está bem compreendida. A alguma distância em ambos os lados do locus *MAT* existem segmentos de DNA chamados *cassetes de silenciamento de tipo reprodutivo* que contêm informação genética extra do tipo reprodutivo (Fig. 21.18). O cassete à esquerda, chamado *HML $\alpha$* , contém informações para o tipo  $\alpha$  de reprodução e o cassete à direita, chamado *HMR<sub>a</sub>*, contém informações para o tipo reprodutivo “a”. Entretanto, um mecanismo de silenciamento gênico impede que cada um desses cassetes seja expresso. Assim, o tipo reprodutivo de uma célula é controlado exclusivamente pela informação genética do locus *MAT*. Quando uma célula de levedura expressa uma enzima chamada HO-endonuclease, o DNA no locus *MAT* pode ser substituído pelo DNA de um dos cassetes flanqueadores. Se o locus *MAT* leva um alelo *a*, o alelo silencioso  $\alpha$  situado no cassete a sua esquerda pode ser copiado no locus *MAT* e subsequente expresso. Essa substituição de DNA fará com que o tipo reprodutivo da célula mude de “a” para  $\alpha$ . Contrariamente, se o locus *MAT* leva o alelo  $\alpha$ , o alelo silencioso *a* no cassete a sua direita pode ser copiado no locus *MAT* e então expresso, mudando o tipo reprodutivo de  $\alpha$  para “a”.

Uma característica importante de todo o sistema de tipo reprodutivo de levedura é que uma célula particular de levedura expressa a informação genética para apenas um dos dois tipos reprodutivos possíveis. Assim, os cassetes flanqueadores do locus expresso *MAT* devem ser silenciados. Pesquisas recentes demonstraram que o silenciamento desses cassetes envolve várias proteínas, incluindo uma chamada de proteína repressor/ativador, abreviada por Rap1p; várias proteínas reguladoras de informação silenciosas, abreviadas por Sir(n)p,





**Fig. 21.18** ■ O processo de mudança de tipo reprodutivo em levedura.

onde  $n$  identifica a proteína específica; e várias proteínas no complexo de reconhecimento da origem, abreviadas por ORC, um complexo que também está envolvido em replicação do DNA. Evidências experimentais indicam que a histona H4 está envolvida no silenciamento dos cassetes de tipo reprodutivo da esquerda e da direita. Embora o mecanismo detalhado ainda não tenha sido elucidado, parece que as proteínas ORC e Rap1p interagem com seqüências específicas de DNA situadas ao longo dos *loci* de tipos reprodutivos em cassetes à esquerda e à direita. Tais proteínas ligadas então facilitam a formação de um complexo protéico mais elaborado que inclui as proteínas Sir, que interagem com Rap1p, histonas e outras para silenciar os genes nos cassetes de tipo reprodutivo à esquerda e à direita. As seqüências de DNA que medeiam este processo são, portanto, apropriadamente chamadas de **silenciadoras**. Quando esse grande complexo protéico foi montado ao redor de seqüências silenciadoras, ele torna os cassetes de tipo reprodutivo da esquerda e da direita inacessíveis a enzimas que catalisam a transição. Silenciamento também foi observado em telômeros de cromossomos de levedura, onde podem controlar a atividade de alguns tipos de elementos de transposição.

### Silenciamento dos Genes *vsg* em Tripanossomas

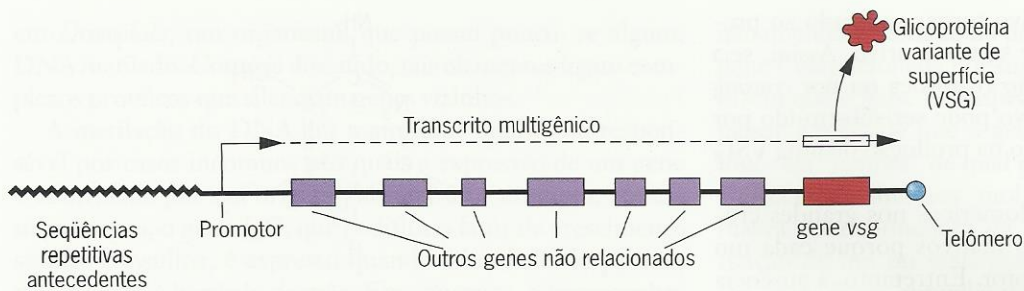
Nos tripanossomas que causam nagana e a doença africana do sono, uma forma de silenciamento controla a expressão dos genes da glicoproteína variante de superfície (*vsg*) mencionada no começo deste capítulo. Estes genes estão situados em vários locais no genoma do tripanossoma, a maioria deles estando perto das pontas dos cromossomos. Pode haver até 1.000 genes *vsg* diferentes. Entretanto, no sangue de um animal infectado, os tripanossomas expressam apenas

um dos genes *vsg*, e ele está sempre situado perto da ponta de um cromossomo. Todos os outros genes *vsg*, incluindo todos os genes *vsg* que estão situados perto de telômeros, estão silenciados. O único gene *vsg* expresso reside em uma unidade multigênica controlada por um promotor distante cerca de 40 a 50 kb (Fig. 21.19). Outros genes não relacionados ficam entre o promotor e o gene *vsg*. Toda esta unidade multigênica é chamada de **sítio de expressão (ES)**.

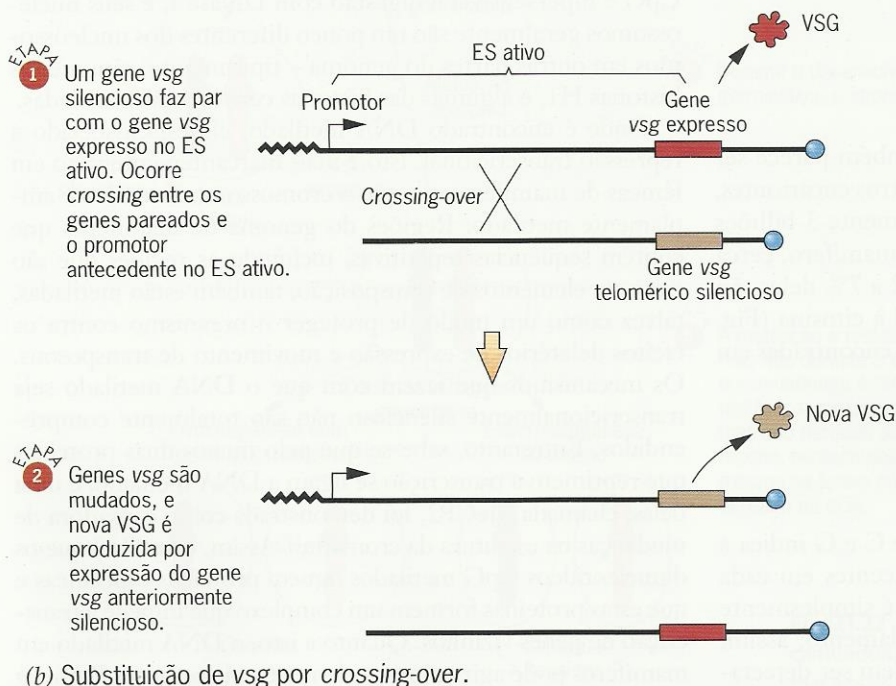
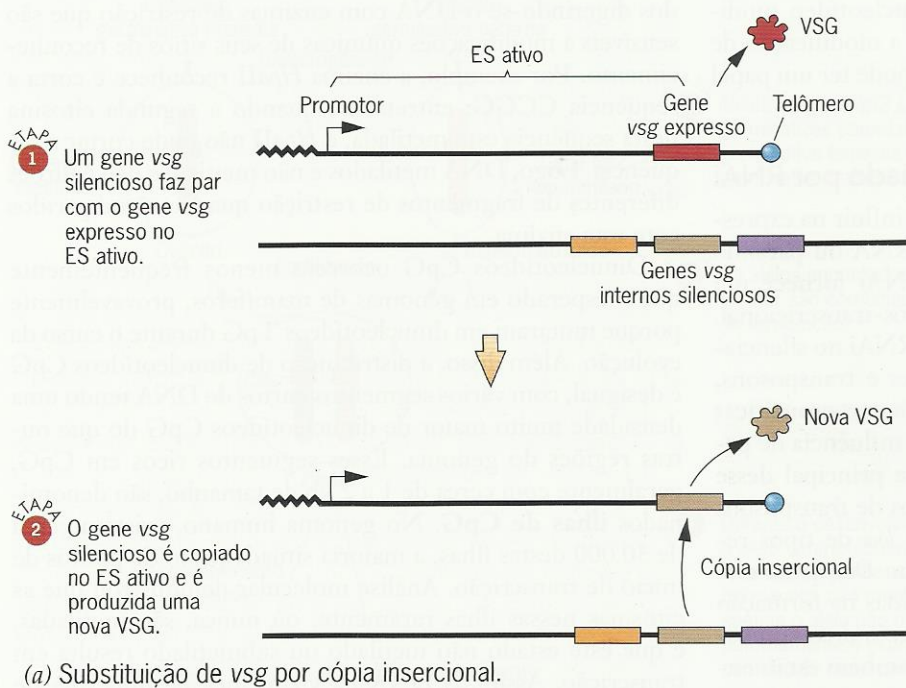
Sítios de expressão parecem existir nas pontas de cada um dos 24 grandes cromossomos de tripanossoma. Entretanto, há qualquer tempo, apenas um deles é transcrito. Um mecanismo ainda desconhecido silencia os outros potenciais sítios de expressão, restringindo assim os tripanossomas infectantes à produção de um só tipo de glicoproteína variante de superfície (VSG). Durante o curso de uma infecção, o ES que é transcrito pode mudar. O ES ativo pode ser silenciado, e um ES antes inativo por tornar-se ativo, resultando na produção de uma VSG diferente. Tais mudanças fornecem um meio para os tripanossomas alterarem suas capas de superfície e escaparem de destruição pelo sistema imunológico do animal infectado.

A capa de superfície do tripanossoma pode ser alterada de dois outros modos. Além dos genes *vsg* situados nos telômeros dos cromossomos grandes, os tripanossomas possuem vários outros genes *vsg* situados em *loci* internos e nos telômeros de mais de 100 minicromossomos. Tais genes *vsg* extras estão todos transcricionalmente silenciados. Entretanto, como nos cassetes de silenciamento do tipo reprodutivo de levedura, um dos genes *vsg* silenciosos pode ser copiado no ES que atualmente está sendo transcrito, fazendo assim com que o organismo produza uma glicoproteína variante de superfície diferente (Fig. 21.20a). A natureza da proteína VSG também pode ser alterada por um *crossing* que troca o gene ativo *vsg* por um dos muitos genes *vsg* teloméricos inativos





**Fig. 21.19** ■ Estrutura de um sítio de expressão (ES) contendo um gene *vsg* ativo perto da ponta de um dos cromossomos em *Trypanosoma brucei*.



**Fig. 21.20** ■ (a) Substituição do gene *vsg* telomérico em um ES ativo de tripanossoma por um gene *vsg* silencioso que codifica uma variante diferente de glicoproteína de superfície (VSG). O mecanismo de substituição remove o gene *vsg* atualmente ativo e insere uma cópia do gene *vsg* silencioso no sítio vago dentro do ES ativo. No final deste processo, é produzida uma nova glicoproteína de superfície variante. (b) Substituição do gene *vsg* telomérico em um ES ativo de tripanossoma por um gene *vsg* telomérico silencioso codificando uma VSG diferente. O mecanismo de substituição envolve *crossing-over* entre os genes *vsg* pareados e o promotor do ES.



(Fig. 21.20b). Após o gene inativo tornar-se ligado ao promotor de um ES transcrito, ele se torna ativo. Assim, seja por um mecanismo de cópia não recíproca ou por *crossing* recíproco, o gene *vsg* no ES ativo pode ser substituído por um gene *vsg* diferente, resultando na produção de uma VSG diferente.

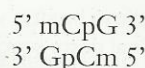
Os genes *vsg* fora dos ES teloméricos nos grandes cromossomos do tripanossoma são inativos porque cada um desses genes não tem um promotor. Entretanto, a ausência de um promotor não pode explicar por que todos menos um dos genes *vsg* situados nos ES teloméricos são inativos. Recentemente, pesquisadores descobriram que os ES inativos contêm  $\beta$ -glicosilidroximetiluracil, um nucleotídeo modificado, ausente no único ES ativo. Assim, a modificação de seqüências de nucleotídeos dentro dos ES pode ter um papel no silenciamento gênico de *vsg* telomérico.

### Silenciamento de Transcrição Mediado por RNAi

Já vimos que o RNA de interferência pode influir na expressão gênica causando a degradação dos mRNA ou parando a tradução destas moléculas. Assim, o RNAi fornece um mecanismo para o silenciamento gênico pós-transcricional. Pesquisas recentes também envolveram o RNAi no silenciamento transcricional de genes. Transgenes e transposons, especialmente aqueles situados em regiões heterocromáticas de um genoma, podem ser silenciados pela influência de pequenos RNA de interferência. A evidência principal desse silenciamento transcricional veio de estudos de transposons em *C. elegans*, transgenes em *Drosophila* e loci de tipos reprodutivos em *Schizosaccharomyces pombe*. Em *Drosophila* e *S. pombe*, vias de RNAi parecem estar envolvidas na formação de heterocromatina ao redor dos centrômeros dos cromossomos. Entretanto, muitos outros fatores também estão envolvidos, e eles podem, de fato, ter papéis importantes nesta forma de silenciamento transcricional.

### Metilação de DNA e Imprinting

A modificação química de nucleotídeos também parece ser importante para a regulação de genes em outros eucariotes, especialmente mamíferos. Dos aproximadamente 3 bilhões de pares de bases em um genoma típico de mamífero, cerca de 40% são pares de bases G:C e cerca de 2 a 7% deles são modificados pela adição de um grupo metil à citosina (Fig. 21.21). A maioria das citosinas metiladas são encontradas em duplas de pares de bases com a estrutura.



onde mC indica metilcitosina e o p entre C e G indica a ligação fosfodiéster entre nucleotídeos adjacentes em cada filamento de DNA. Esta estrutura em geral é simplesmente abreviada dando-se a composição de um filamento; assim, mCpG. Dinucleotídeos CpG metilados podem ser detecta-

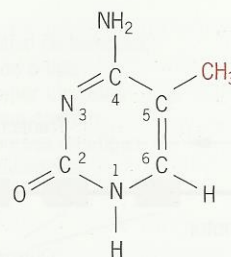


Fig. 21.21 ■ A estrutura de 5-metilcitosina.

dos digerindo-se o DNA com enzimas de restrição que são sensíveis a modificações químicas de seus sítios de reconhecimento. Por exemplo, a enzima *HpaII* reconhece e corta a seqüência CCGG; entretanto, quando a segunda citosina nesta seqüência está metilada, a *HpaII* não pode cortar a seqüência. Logo, DNA metilados e não metilados dão padrões diferentes de fragmentos de restrição quando são digeridos com esta enzima.

Dinucleotídeos CpG ocorrem menos frequentemente que o esperado em genomas de mamíferos, provavelmente porque mutaram em dinucleotídeos TpG durante o curso da evolução. Além disso, a distribuição de dinucleotídeos CpG é desigual, com vários segmentos curtos de DNA tendo uma densidade muito maior de dinucleotídeos CpG do que outras regiões do genoma. Esses segmentos ricos em CpG, geralmente com cerca de 1 a 2 kb de tamanho, são denominados **ilhas de CpG**. No genoma humano, existem cerca de 30.000 destas ilhas, a maioria situada perto de pontos de início de transcrição. Análise molecular demonstrou que as citosinas nessas ilhas raramente, ou nunca, são metiladas, e que este estado não metilado ou submetilado resulta em transcrição. Assim, o DNA na vizinhança de uma ilha de CpG é hipersensível a digestão com DNase I, e seus nucleossomos geralmente são um pouco diferentes dos nucleossomos em outras partes do genoma – tipicamente, têm menos histonas H1, e algumas das histonas cernas estão acetiladas.

Onde é encontrado DNA metilado, ele está associado a repressão transcricional. Isto é mais marcadamente visto em fêmeas de mamíferos em que o cromossomo X inativo é amplamente metilado. Regiões do genoma de mamíferos que contêm seqüências repetitivas, incluindo as regiões que são ricas em elementos de transposição, também estão metiladas, talvez como um modo de proteger o organismo contra os efeitos deletérios de expressão e movimento de transposons. Os mecanismos que fazem com que o DNA metilado seja transcricionalmente silencioso não são totalmente compreendidos. Entretanto, sabe-se que pelo menos duas proteínas que reprimem a transcrição se ligam a DNA metilado, e uma delas, chamada MeCP2, foi demonstrada como causadora de mudanças na estrutura da cromatina. Assim, é possível que os dinucleotídeos CpG metilados liguem proteínas específicas e que estas proteínas formem um complexo que impede a transcrição de genes vizinhos. Quanto a isto, o DNA metilado em mamíferos pode agir como os elementos de resposta *Polycomb*

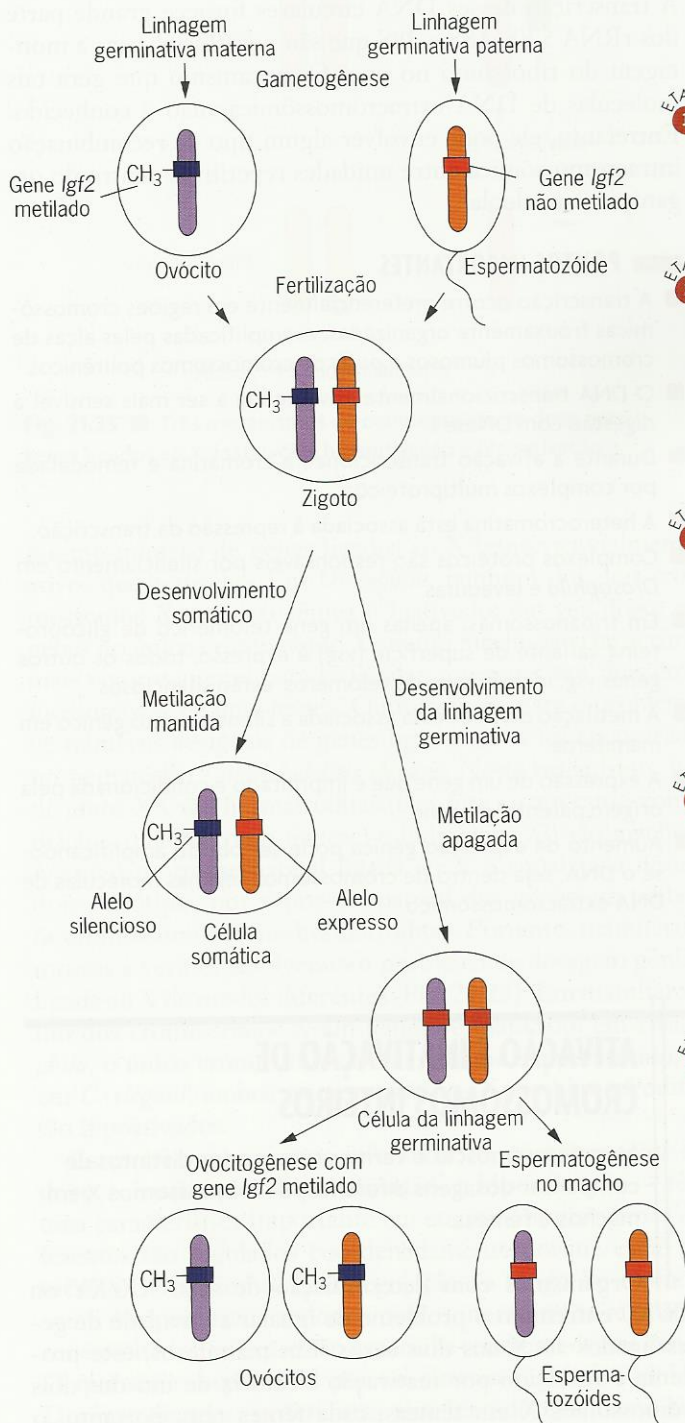


em *Drosophila*, um organismo que possui pouco, se algum, DNA metilado. Como já discutido, tais elementos ligam complexos protéicos que silenciam genes vizinhos.

A metilação do DNA em mamíferos também é responsável por casos incomuns nos quais a expressão de um gene é controlada por sua origem parental. Por exemplo, em camundongos, o gene *Igf2*, que codifica o fator de crescimento similar à insulina, é expresso quando é herdado do pai mas não quando é herdado da mãe. Em contraste, o gene conhecido como *H19* é expresso quando é herdado da mãe mas

não quando é herdado do pai. Sempre que a expressão de um gene é condicionada por sua origem parental, os geneticistas dizem que o gene foi **imprintado** – um termo que pretende passar a idéia de que o gene foi marcado de algum modo; logo, ele “lembra” de qual genitor ele veio.

Recentes análises moleculares demonstraram que a marca condicionadora da expressão de um gene é a metilação de um ou mais dinucleotídeos CpG na vizinhança do gene. Esses dinucleotídeos metilados são inicialmente formados na linhagem germinativa parental (Fig. 21.22).



**Fig. 21.22** ■ Metilação e imprinting do gene *Igf2* em camundongos. O gene é metilado em fêmeas mas não em machos.



Assim, por exemplo, o gene *Igf2* é metilado na linhagem germinativa feminina mas não na linhagem germinativa masculina. Na fertilização, um gene metilado *Igf2* contribuído pela mãe é combinado com um gene *Igf2* não metilado contribuído pelo pai. Durante a embriogênese, os estados metilado e não metilado são preservados cada vez que os genes se replicam. Como um gene metilado é silencioso, apenas o gene *Igf2* contribuído paternamente é expresso no animal em desenvolvimento. Exatamente o oposto ocorre com o gene *H19*, que é metilado na linhagem germinativa masculina mas não na linhagem germinativa feminina. Mais de 20 genes imprintados diferentes foram identificados em camundongos e seres humanos. Para cada um, o *imprint* de metilação é estabelecido na linhagem germinativa parental. Entretanto, um gene metilado que foi herdado de um sexo pode ser desmetilado quando passa por uma prole do sexo oposto. Assim, os *imprints* de metilação são refeitos a cada geração, dependendo do sexo do animal. O fato de que alguns genes são metilados em um sexo mas não no outro significa que fatores específicos do sexo controlam a maquinaria de metilação.

## Amplificação Gênica

Às vezes, a expressão gênica é facilitada por um aumento no número de genes. Este processo, chamado **amplificação gênica**, é destinado a aumentar o número de moldes de DNA para síntese de RNA. Talvez o exemplo mais marcante de amplificação gênica relacionada a expressão é o que envolve os genes de RNA ribossômico em ovócitos de anfíbios. Estes genes são necessários para produzir componentes estruturais dos ribossomos. Em eucariotes, existem quatro tipos principais de genes de rRNA: o gene 5S, que codifica um rRNA com 120 bases; o gene 5,8S, que codifica um rRNA com 160 bases; o gene 18S, que codifica um rRNA com 1,8 kb; e o gene 28S, que codifica um rRNA com 4,7 kb. No genoma do sapo *Xenopus laevis*, milhares de genes 5S estão distribuídos por todos os cromossomos, mas existem bem menos cópias dos genes de 5,8S, 18S e 28S, talvez apenas 800 a 1.000 por célula diplóide. Estes últimos genes estão todos concentrados em um único sítio denominado **organizador nucleolar**, assim chamado porque forma o nucléolo. O mapeamento molecular mostrou que os genes de rRNA 5,8S, 18S e 28S pertencem a uma única unidade transcricional que é repetida em tandem dentro do organizador nucleolar.

Todos os quatro tipos de genes de RNA ribossômico devem ser transcritos para produzir rRNA para ribossomos de *Xenopus*. Em ovócitos, essa transcrição deve ser especialmente vigorosa porque o ovo precisa acumular ribossomos suficientes para suportar o desenvolvimento inicial do embrião após fertilização. Talvez até  $10^{12}$  ribossomos devam ser produzidos em cada ovócito. Como cada ribossomo contém exatamente uma molécula de cada tipo de rRNA,  $10^{12}$  moléculas devem ser produzidas de cada um dos rRNAs 5S, 5,8S, 18S e 28S. Essa necessidade é atendida de dois modos.

Primeiro, *Xenopus* tem cerca de 24.000 genes 5S, incluindo muitos que são especificamente ativados em ovócitos. Este imenso conjunto de genes é portanto capaz de gerar rRNAs 5S necessários para a produção de ribossomos no ovo. Segundo, o pequeno número de genes de 5,8S, 18S e 28S é especificamente amplificado em ovócitos pela criação de cópias extracromossômicas destes genes. São formados pequenos DNA circulares covalentemente fechados levando genes de 5,8S, 18S e 28S. Eles se replicam pelo mecanismo de círculo rolante (Cap. 10) para produzir muitas cópias, que formam nucléolos supernumerários dentro do ovócito. A transcrição desses DNA circulares fornece grande parte dos rRNA 5,8S, 18S e 28S que são necessários para a montagem do ribossomo no ovo. O mecanismo que gera tais moléculas de DNA extracromossômicas não é conhecido. Entretanto, ele pode envolver algum tipo de recombinação intracromossômica entre unidades repetitivas dentro do organizador nucleolar.

## PONTOS IMPORTANTES

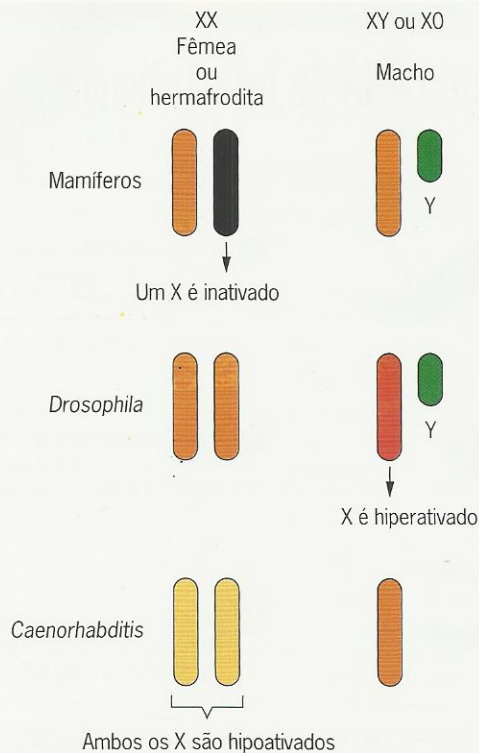
- A transcrição ocorre preferencialmente em regiões cromossômicas frouxamente organizadas, exemplificadas pelas alças de cromossomos plumosos e *puffs* de cromossomos politênicos.
- O DNA transcricionalmente ativo tende a ser mais sensível a digestão com DNase I.
- Durante a ativação transcricional, a cromatina é remodelada por complexos multiprotéicos.
- A heterocromatina está associada à repressão da transcrição.
- Complexos protéicos são responsáveis por silenciamento em *Drosophila* e leveduras.
- Em tripanossomas, apenas um gene telomérico de glicoproteína variante de superfície (*vsg*) é expresso; todos os outros genes *vsg*, incluindo os de telômeros, estão silenciosos.
- A metilação do DNA está associada a silenciamento gênico em mamíferos.
- A expressão de um gene que é imprintado é condicionada pela origem parental do gene.
- Aumento da expressão gênica pode ser obtido amplificando-se o DNA, seja dentro de cromossomos seja nas moléculas de DNA extracromossômico.

## ATIVAÇÃO E INATIVAÇÃO DE CROMOSSOMOS INTEIROS

Mamíferos, moscas e vermes têm modos distintos de compensar dosagens diferentes de cromossomos X em machos e fêmeas.

Organismos com determinação de sexo XX/XY ou XX/X0 enfrentam o problema de igualar a atividade de genes ligados ao X nos dois sexos. Em mamíferos, este problema é resolvido por inativação aleatória de um dos dois cromossomos X em fêmeas; cada fêmea tem, portanto, o





**Fig. 21.23** ■ Três mecanismos de compensação de dose para genes ligados ao X: inativação, hiperativação e hipoativação.

mesmo número de genes ligados ao X transcricionalmente ativos que o macho. Em *Drosophila*, nenhum dos dois cromossomos X em uma fêmea é inativado; em vez disso, os genes no único cromossomo X em um macho são transcritos mais vigorosamente para se igualar aos genes nos dois cromossomos X de uma fêmea. Outra solução para o problema de números desiguais de genes ligados ao X foi encontrada no nematódeo *Caenorhabditis elegans*. Neste organismo, indivíduos XX são hermafroditas (eles funcionam tanto como machos quanto como fêmeas) e indivíduos XO são machos. A atividade transcricional ligada ao X é equalizada nestes dois genótipos por repressão parcial dos genes em ambos os cromossomos X nos hermafroditas. Portanto, mamíferos, moscas e vermes resolveram o problema de dosagem gênica ligada ao X de modos diferentes (Fig. 21.23). Em mamíferos, um dos cromossomos X em fêmeas é inativado; em *Drosophila*, o único cromossomo X em machos é hiperativado; e, em *C. elegans*, ambos os cromossomos X em hermafroditas são hipoativados.

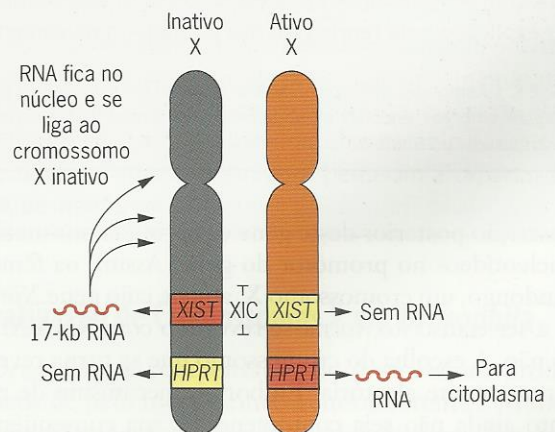
Estes três mecanismos diferentes de **compensação de dose** – **inativação**, **hiperativação** e **hipoativação** – têm uma característica importante em comum: muitos genes diferentes são regulados coordenadamente porque estão no mesmo cromossomo. Essa regulação ao longo do cromossomo é superposta a todos os outros mecanismos reguladores envolvidos na expressão espacial e temporal desses genes. O que seria responsável por tal sistema regulador global? Durante décadas, geneticistas têm tentado elucidar a base

molecular da compensação de dose. A hipótese de trabalho tem sido que algum fator (ou fatores) se liga especificamente ao cromossomo X e altera suas atividades transcricionais. Descobertas recentes indicam que esta idéia é correta.

## Inativação de Cromossomos X em Mamíferos

Em mamíferos, a inativação do cromossomo X começa em um local particular chamado **centro de inativação do X (XIC)** e então se espalha em sentidos opostos até as pontas do cromossomo. Curiosamente, nem todos os genes em um cromossomo X inativado estão transcricionalmente silenciosos. Um dos que permanecem ativos é denominado *XIST* (de transcrito específico do X inativo). Este gene está situado dentro de XIC (Fig. 21.24). Pesquisa recente mostrou que, em seres humanos, o gene *XIST* codifica um transcrito de 17 kb desprovido de matrizes de leitura aberta significativas. Portanto, parece improvável que o gene *XIST* codifique uma proteína. Em vez disso, o próprio RNA provavelmente é o produto funcional do gene *XIST*. Embora poliadenilado, esse RNA é restrito ao núcleo e está especificamente situado nos cromossomos X inativados. Ele não parece estar associado aos cromossomos X ativos nem em machos nem em fêmeas.

Em camundongos, nos quais algumas análises experimentais foram possíveis, pesquisadores descobriram que o homólogo do gene humano *XIST* é transcrito durante os primeiros estágios do desenvolvimento embrionário em um nível baixo por ambos os cromossomos X que estão presentes em fêmeas. Os transcritos de cada um dos genes *Xist* de uma fêmea de camundongo são instáveis e permanecem bem associados a seus respectivos genes. À medida que o desenvolvimento continua, os transcritos de um dos genes se estabiliza e envolve todo o cromossomo X no qual o gene está situado. Os transcritos do outro gene *Xist* desintegram-se, e



**Fig. 21.24** ■ Expressão do gene *XIST* no cromossomo X inativo de mulheres. É mostrada, para comparação, a expressão do gene *HPRT* no cromossomo X ativo. Este gene codifica hipoxantina-fosforibosiltransferase, uma enzima que desempenha um papel no metabolismo de purinas.



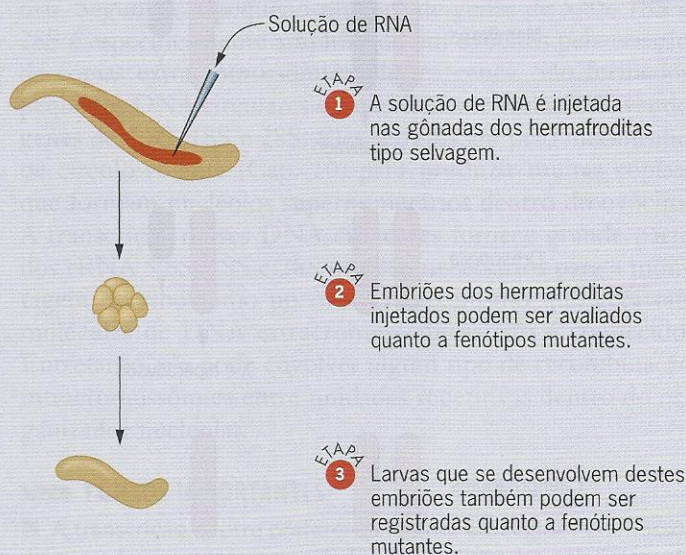


## UM MARCO NA GENÉTICA: A Descoberta do RNA de Interferência

Às vezes, uma descoberta científica emerge de um esforço para explicar um paradoxo. A descoberta de que o RNA bifilamentar induz uma via que interfere na expressão gênica é um destes casos. Em 1995, Su Guo e Kenneth Kemphues publicaram uma análise de *par-1*, um gene de ação materna de *C. elegans*.<sup>1</sup> O produto deste gene ajuda a estabelecer a polaridade anterior-posterior em embriões de *C. elegans*. Hermafroditas homozigotos para uma mutação nula *par-1* produzem embriões que param no primeiro estágio do desenvolvimento. Esses embriões têm muitas células musculares na faringe e não têm células intestinais. Consequentemente, eles morrem. Guo e Kemphues isolaram um cDNA, chamado ZC22, que pensaram ter vindo do gene *par-1*. Para testar a autenticidade deste cDNA, eles injetaram RNA anti-sentido feito de ZC22 nas gônadas de hermafroditas tipo selvagem e avaliaram os embriões resultantes quanto ao fenótipo mutante *par-1* (Fig. 1).

O experimento foi um sucesso (veja Quadro 1). O RNA anti-sentido feito de ZC22 fez com que os embriões parecessem ter vindo de genitores homozigotos para *par-1*. Esses embriões tinham muitas células musculares faringiais, não tinham células intestinais e 52% deles morriam. Guo e Kemphues não observaram o fenótipo mutante *par-1* quando injetaram hermafroditas com uma solução de água ou quando os injetaram com soluções de RNA anti-sentido feito de outros genes. Assim, eles estavam confiantes de que o cDNA que eles haviam testado vinha do gene *par-1*.

Guo e Kemphues fizeram mais um experimento controle. Eles injetaram RNA com sentido feito de ZC22 em hermafroditas tipo selvagem e examinaram os embriões resultantes. Para sua surpresa, eles observaram que esses embriões tinham o fenótipo mutante *par-1* e que 54% deles morriam. Assim, procurando certificar-se, Guo e Kemphues esbarraram em um paradoxo. Eles descobriram que as injeções de RNA anti-sentido ZC22 produziam o fenótipo mutante *par-1*, como esperado. Mas assim também as injeções de RNA com sentido de ZC22. O efeito do RNA anti-sentido podia ser explicado pela teoria genética padrão – o pareamento de



**Fig. 1** ■ Esquema para testar o RNA quanto à capacidade de interferir na expressão de um gene tipo selvagem em *C. elegans*. O efeito fenotípico da solução teste é analisado em embriões ou larvas que se desenvolvem dos animais injetados.

bases entre o RNA anti-sentido injetado e o RNA mensageiro de *par-1* bloquearia a tradução do RNA mensageiro de *par-1* em uma proteína –, mas o efeito do RNA com sentido não podia ser prontamente explicado.

Em 1998, pesquisadores descobriram uma explicação para o efeito paradoxal do RNA com sentido. Andrew Fire, SiQun Xu, Mary Montgomery, Steven Kostas, Samuel Driver e Craig Mello publicaram um trabalho descrevendo os efeitos de preparações de RNA uni- e bifilamentares na expressão de genes diferentes de *C. elegans*.<sup>2</sup> Um grupo de experimentos enfocou o *unc-22*, um

<sup>1</sup>Guo, S., e K. J. Kemphues. 1995. *par-1*, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell* 81:611-620.

<sup>2</sup>Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver e C. C. Mello. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391: 806-811.

a transcrição posterior deste gene é reprimida por metilação de nucleotídeos no promotor do gene. Assim, na fêmea de camundongo, um cromossomo X, aquele cujo gene *Xist* continua a ser transcrito, torna-se revestido com RNA *Xist* e o outro não. A escolha do cromossomo que se torna revestido é aparentemente aleatória. Embora o mecanismo de revestimento ainda não seja compreendido, sua consequência é clara: a maioria dos genes no cromossomo revestido estão reprimidos, e esse cromossomo torna-se o X inativo. No sistema de compensação de dose de mamíferos, portanto, o cromossomo X que permanece ativo é, paradoxalmente, o que reprime seu gene *Xist*.

Cromossomos X inativos são prontamente identificados em células de mamíferos. Durante a intérfase, eles se condensam em uma massa de coloração escura associada à membrana nuclear. Essa massa, o corpúsculo de Barr, descondensa-se durante a fase S para permitir que o X inativo seja replicado. Entretanto, como a descondensação leva algum tempo, o X inativo replica-se depois do resto dos cromossomos. Os cromossomos X inativos devem portanto ter uma estrutura de cromatina muito diferente da dos outros cromossomos. Esta diferença é parcialmente determinada pelos tipos de histonas associadas ao DNA. Uma das quatro histonas cernas, a H4, pode ser modificada quimicamente





## UM MARCO NA GENÉTICA (continuação)

**Quadro 1** Fenótipos Mutante (*par-1*) e Tipo Selvagem em Embriões de *C. elegans* Hermafroditas Injetados por Guo e Kemphues

| Injeção da Solução                                        | Fenótipo Embrionário |
|-----------------------------------------------------------|----------------------|
| RNA anti-sentido ZC22                                     | mutante              |
| Água                                                      | tipo selvagem        |
| RNA anti-sentido TS<br>(de um gene de <i>Drosophila</i> ) | tipo selvagem        |
| RNA ZC22 com sentido                                      | mutante              |

**Quadro 2** Fenótipos Mutante (*unc-22*) e Tipo Selvagem em Larvas de *C. elegans* Hermafroditas Tipo Selvagem Injetadas por Fire e Colaboradores

| Injeção da Solução                  | Fenótipo Larval |
|-------------------------------------|-----------------|
| RNA anti-sentido de <i>unc-22</i>   | tipo selvagem   |
| RNA com sentido de <i>unc-22</i>    | tipo selvagem   |
| RNA + anti-sentido de <i>unc-22</i> | mutante         |
| RNA com sentido de <i>unc-22</i>    |                 |

gene que codifica uma proteína do miofilamento encontrada em células musculares (veja Quadro 2). Quando a expressão de *unc-22* é prejudicada, *C. elegans* exibe um comportamento contorcido. Como este fenótipo é facilmente registrado, ele fornece uma análise conveniente para detectar o efeito do RNA injetado na expressão de *unc-22*. Fire e seus colegas procuraram o fenótipo contorcido na prole de vermes do tipo selvagem que tinham sido injetados com RNA unifilamentar, seja com sentido seja com anti-sentido, ou bifilamentar do gene *unc-22*. O DNA bifilamentar foi obtido misturando-se os RNA com sentido e anti-sentido.

Os resultados destes experimentos foram definitivos. Nenhum dos tipos de RNA unifilamentar, com sentido ou anti-sentido, teve um efeito significativo sobre o fenótipo da prole dos vermes injetados. Entretanto, o RNA bifilamentar fez com que esta prole fosse contorcida, isto é, mimetizou o fenótipo de um mutante

*unc-22*. O RNA de *unc-22* bifilamentar injetado portanto interferiu na expressão do gene endógeno *unc-22*.

Trabalhos subsequentes mostraram que este tipo de interferência envolve a enzima Dicer, que corta o RNA bifilamentar injetado em pequenos segmentos, e um complexo de riboproteínas chamado RISC, que usa tais segmentos como guias para ter como alvo RNA mensageiros complementares para degradação. Os experimentos de Fire e colaboradores, que foram originalmente destinados a resolver o paradoxo encontrado por Guo e Kemphues, descobriram um novo fenômeno: o RNA de interferência.

Mas por que o RNA com sentido que Guo e Kemphues injetaram prejudica a expressão gênica? Mais provavelmente, não foi o puro RNA com sentido, mas uma mistura de moléculas com sentido e anti-sentido. O RNA bifilamentar é um indutor altamente eficiente de interferência de RNA. Mesmo algumas moléculas bifilamentares irão disparar a via de interferência. Se o RNA com sentido que Guo e Kemphues usaram estivesse contaminado com RNA anti-sentido, algumas moléculas de RNA bilamentar teriam sido formadas na solução. Vermes injetados com esta solução responderiam apropriadamente direcionando o RNA mensageiro endógeno para degradação. O RNA com sentido pareceria ter interferido na expressão do gene alvo. O efeito paradoxal do RNA com sentido, portanto, mais provavelmente foi causado por contaminantes anti-sentido na solução de injeção.

### QUESTÕES PARA DISCUSSÕES

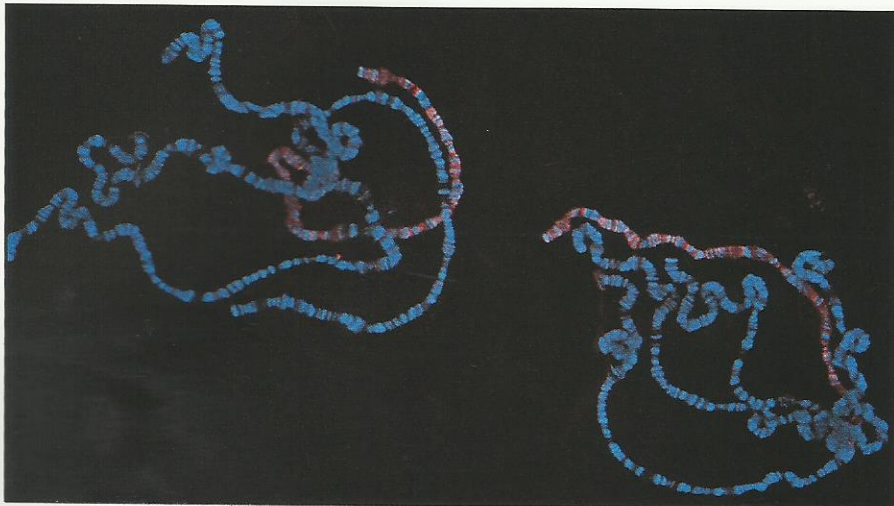
1. Como o RNA bifilamentar foi usado para estudar a função gênica em organismos modelos tais como *C. elegans* e *Drosophila*? Como foi usado para estudar o funcionamento gênico em organismos que não são passíveis de análise genética padrão? O RNA bifilamentar tem o potencial de agente terapêutico para atenuar ou curar doenças humanas?
2. O RNA de interferência tornou-se um instrumento útil na pesquisa genética. Entretanto, ele já existia muito antes de geneticistas o descobrirem. Por que organismos têm vias de RNA de interferência? Quando podem ter evoluído as vias de RNA de interferência? Alguns organismos perderam a capacidade de RNA de interferência?

pela adição de grupos acetil a qualquer das várias lisinas na cadeia polipeptídica. A H4 acetilada está associada a todos os cromossomos no genoma humano. Entretanto, no X inativo ela parece estar restrita a três bandas um tanto estreitas, cada uma correspondendo a uma região que contém alguns genes ativos. A H4 acetilada também está ausente em áreas de heterocromatina nos outros cromossomos. Estes achados sugerem que a falta de H4 acetilada é uma característica fundamental do cromossomo X inativo.

### Hiperativação de Cromossomos X em *Drosophila*

Em *Drosophila*, a compensação de dose requer produtos protéicos de pelo menos cinco genes diferentes. Mutações nulas nestes genes resultam em letalidade específica de machos, pois o único cromossomo X em machos não é hiperativado. Machos mutantes geralmente morrem durante os estágios larval e inicial de pupa. Estes genes de compensação de dose são portanto chamados de loci letais específicos de machos (*msl*), e seus produtos são chamados de proteínas





**Fig. 21.25** ■ Ligação do produto protéico do gene *mle* de *Drosophila* ao único cromossomo X em machos.

MSL. Anticorpos preparados contra tais proteínas foram usados como sondas para localizar as proteínas dentro de células. O achado marcante é que cada uma das proteínas MSL se liga especificamente ao cromossomo X em machos (Fig. 21.25). Uma dessas proteínas, o produto do gene *mle* chamado *maleless* (*mle*), é homóloga às helicases de DNA e RNA, que são enzimas capazes de deselucidar ácidos nucleicos. A suposta função da helicase da proteína Maleless é consistente com a idéia de que o cromossomo X em uma *Drosophila* deve estar “aberto” para uma intensa transcrição. Ela também é compatível com a observação de que cromossomos X politênicos em machos parecem inchados e têm um padrão difuso de bandamento, talvez porque estão hiperativados.

As cinco proteínas MSL aparentemente formam um complexo que se liga ao único cromossomo X em machos de *Drosophila*. Estudos recentes indicam que esse complexo de proteínas MSL também contém tipos específicos de RNA. Dois RNA diferentes, *roX1* e *roX2* (de RNA no cromossomo X), estão co-localizados no cromossomo X do macho com as proteínas MSL. Ambos os RNA são transcritos de genes ligados ao X, e ambos são recompostos e poliadenilados. Entretanto, eles não contêm longas matrizes de leitura abertas, e portanto não são traduzidos em proteínas. Como os RNA *Xist* de mamíferos, os RNA *roX* em *Drosophila* parecem ser produtos finais funcionais de genes que os codificam. Curiosamente, a expressão dos genes *roX* depende do funcionamento apropriado dos genes *mle*. Além disso, esta expressão é desligada quando *Sxl*, o principal gene regulador de compensação de dose de *Drosophila* e sistemas de determinação do sexo, é desligado, como naturalmente é em fêmeas de *Drosophila*. Os RNA de *roX* são portanto produzidos apenas em machos de *Drosophila*.

O modelo atual propõe que as proteínas MSL formam um complexo que é unido pelos RNA de *roX*. Este complexo então se une a 30 a 40 sítios ao longo do cromossomo X do macho, incluindo os *loci* que contêm os dois genes *roX*.

A partir de cada um desses sítios de entrada, o complexo MSL/*roX* prolonga-se bidirecionalmente até atingir todos os genes no cromossomo X do macho que precisam ser hiperativados. O processo de hiperativação pode envolver remodelagem da cromatina pelo complexo MSL/*roX*. Uma das proteínas MSL é uma histona-acetiltransferase, e uma versão particular da histona H4 está exclusivamente associada a cromossomos X hiperativados.

### Hipoativação de Cromossomos X em *Caenorhabditis*

Em *C. elegans*, a compensação de dose envolve a repressão parcial de genes ligados ao X nas células somáticas de hermafroditas. O mecanismo não é totalmente compreendido, mas os produtos de vários genes estão envolvidos. Como as proteínas MSL em *Drosophila*, as proteínas codificadas por estes genes ligam-se especificamente ao cromossomo X. Entretanto, ao contrário da situação em *Drosophila*, elas se ligam apenas quando estão presentes dois cromossomos X. As proteínas aparentemente não se ligam ao único cromossomo X em machos nem se ligam a nenhum dos autossomos tanto em machos quanto em hermafroditas. A compensação de dose em *C. elegans* parece portanto envolver um mecanismo exatamente oposto ao de *Drosophila*. Um complexo protéico liga-se aos cromossomos X e reprime, em vez de intensificar, a transcrição.

#### PONTOS IMPORTANTES

- A inativação de um cromossomo X em fêmeas XX de mamíferos é mediada por um RNA não codificante transcrito do gene *XIST* nesse cromossomo.
- A hiperativação do único cromossomo X no macho de *Drosophila* é mediada por um complexo RNA-proteína que se liga a muitos sítios nesse cromossomo e estimula a transcrição de seus genes.
- A hipoinativação dos dois cromossomos X em hermafroditas de *C. elegans* é mediada por proteínas que se ligam a esses cromossomos e reduzem a transcrição de seus genes.



## Exercícios Básicos Ilustram a análise genética básica.

1. Disponha os seguintes eventos em ordem cronológica, começando pelo primeiro: (a) recomposição de uma molécula de RNA, (b) migração de uma molécula de mRNA para o citoplasma, (c) transcrição de um gene, (d) degradação de uma molécula de mRNA, (e) síntese de polipeptídeo.

**Resposta:** c - a - b - e - d.

2. Que fatores induzem a expressão do gene *hsp70* em *Drosophila* e do gene de ribulose-bisfosfato-carboxilase (*RBC*) em plantas?

**Resposta:** O gene *hsp70* é induzido por estresse provocado por calor; o gene *RBC* é induzido por exposição a luz.

3. Indique se cada um dos seguintes fenômenos relacionados à regulação da expressão gênica ocorre no núcleo ou no citoplasma de uma célula eucariótica.

- (a) Estimulação da expressão gênica por um fator de transcrição.
- (b) Recomposição alternativa do transcrito primário de um gene.
- (c) Poliadenilação do transcrito primário de um gene.
- (d) Tradução de um RNA mensageiro.
- (e) Inibição da tradução por um microRNA ligado a um RNA mensageiro.
- (f) Degradação de um RNA mensageiro induzida por um curto RNA de interferência.
- (g) Ligação de um hormônio peptídico a seu receptor.
- (h) Ligação de um hormônio esteróide a seu receptor.
- (i) Silenciamento da expressão gênica por heterocromatina.
- (j) Inativação de um cromossomo inteiro.

**Resposta:** O item (h) pode ocorrer no citoplasma ou no núcleo, dependendo do hormônio esteróide específico. Os itens (a), (b), (c), (i) e (j) ocorrem no núcleo. Todos os outros itens ocorrem no citoplasma.

4. Indique se os seguintes estão associados a atividade gênica ou a inatividade: (a) grupo de proteínas Polycomb, (b) metilação do DNA, (c) acetilação de histona, (d) metilação de histona, (e) heterocromatina, (f) região de controle de *locus*, (g) proteína GAL4, (h) sensibilidade à DNase I.

**Resposta:** (a) Inatividade, (b) inatividade, (c) atividade, (d) inatividade, (e) inatividade, (f) atividade, (g) atividade, (h) atividade.

5. Quais são algumas das diferenças entre eucromatina e heterocromatina?

**Resposta:** A heterocromatina cora-se fortemente durante o ciclo celular; a eucromatina não é intensamente corada durante a interfase. A heterocromatina é rica em seqüências repetidas de DNA e em elementos de transposição; a eucromatina contém seqüências repetidas e transposons, mas geralmente não ao ponto da heterocromatina. A heterocromatina tem alguns genes codificantes de proteínas; a eucromatina tem muitos genes codificantes de proteínas.

6. Como o nível de expressão gênica ligada ao X é igualado nos dois sexos de (a) humanos, (b) moscas, (c) vermes?

**Resposta:** (a) Em humanos, um dos dois cromossomos X em mulheres é aleatoriamente inativado. (b) Em moscas, o cromossomo X frágil em machos é hiperativado. (c) Em vermes, os dois cromossomos X em hermafroditas são hipoativados.

## Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

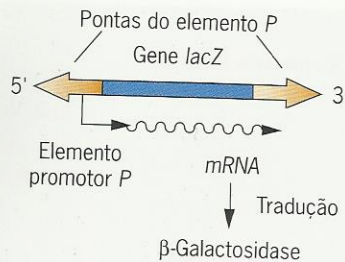
1. O gene bacteriano *lacZ* para  $\beta$ -galactosidase foi inserido em um elemento de transposição *P* de *Drosophila* (Cap. 18) de modo que ele possa ser transcrito a partir do promotor de elemento *P*. Este gene de fusão foi então injetado na linhagem germinativa de um embrião de *Drosophila* juntamente com uma enzima que catalisa a transposição de elementos *P*. Durante o desenvolvimento, um elemento *P* modificado tornou-se inserido nos cromossomos de algumas células da linhagem germinativa. A prole deste animal injetado foi então cruzada individualmente com moscas de um estoque padrão de laboratório para estabelecer linhagens que levam o gene de fusão *P/lacZ* em seus genomas. Três destas linhagens foram analisadas quanto a expressão de *lacZ* corando-se tecidos dissecados de moscas adultas com X-gal, um substrato cromogênico que fica azul na presença de  $\beta$ -galactosidase. Na primeira linhagem, apenas os olhos ficaram azuis; na segunda, apenas os intestinos coraram-se em azul; e, na terceira, todos os tecidos ficaram azuis. Como você explica estes resultados?

**Resposta:** As três linhagens evidentemente levam inserções diferentes do gene de fusão *P/lacZ*. Em cada linhagem, a expressão

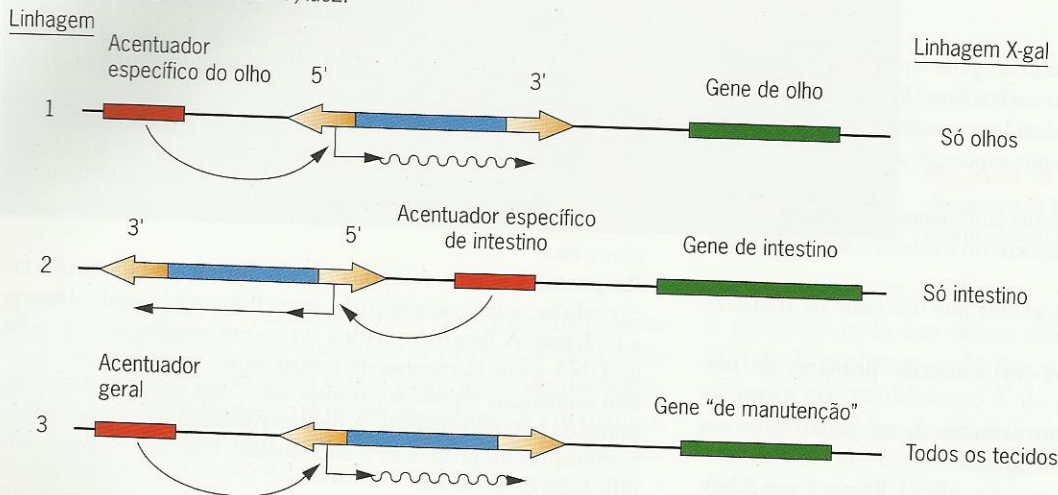
do gene de fusão *P/lacZ* deve ter estado sob a influência de uma seqüência reguladora, ou acentuador, capaz de interagir com o promotor *P* e iniciar a transcrição no gene *lacZ*. Na primeira linhagem, o elemento *P* modificado deve ter-se inserido perto de um acentuador específico de olho, que comandou a transcrição apenas no tecido do olho. Na segunda linhagem, ele deve ter-se inserido perto de um acentuador que comanda a transcrição nas células intestinais. Na terceira linhagem, ele deve ter-se inserido perto de um acentuador que comanda a transcrição em todas, ou quase todas, as células, independente do tipo de tecido. Supostamente, cada um desses acentuadores diferentes está perto de um gene que normalmente seria expresso sob seu controle. Por exemplo, o acentuador específico do olho estaria perto de um gene necessário para algum aspecto de funcionamento ou desenvolvimento do olho. Tais resultados mostram que inserções aleatórias do gene de fusão *P/lacZ* podem ser usadas para identificar tipos diferentes de acentuadores e, através deles, os genes que eles controlam. Essas inserções de genes de fusão são portanto chamadas de *enhancer traps*.



Gene de fusão P/lacZ:



Inserções do gene de fusão P/lacZ:



2. Em sua publicação sobre RNA de interferência, Andrew Fire e colaboradores (1998 *Nature* 391: 806-811) descrevem os resultados de experimentos nos quais o RNA derivado do gene *mex-3* foi injetado em *C. elegans* hermafroditas. Os embriões obtidos desses hermafroditas injetados foram analisados por hibridização *in situ* usando-se sondas para RNA de *mex-3*. As sondas foram projetadas para ligar-se ao RNA mensageiro de *mex-3*, que normalmente se acumula nas gônadas de hermafroditas e em seus embriões. A ligação de moléculas da sonda ao mRNA nos embriões é facilmente detectada se as moléculas da sonda forem marcadas. Quando Fire e seus colaboradores fizeram estes experimentos de hibridização *in situ*, eles descobriram que embriões de vermes que tinham sido injetados com RNA de *mex-3* bifilamentar não foram marcados pelas moléculas da sonda, enquanto embriões que tinham sido injetados com RNA unifilamentar complementar ao mRNA de *mex-3*, isto é, com RNA anti-sentido de *mex-3*, foram marcados embora não tão

intensamente quanto os embriões de vermes que não tinham sido injetados. O que estes resultados indicam sobre a eficácia do RNA anti-sentido bifilamentar *versus* unifilamentar para silenciar a expressão gênica?

**Resposta:** Os resultados destes experimentos de hibridização *in situ* indicam que o RNA bifilamentar é um forte silenciador da expressão do gene *mex-3* em embriões de *C. elegans*. Em contraste, o RNA anti-sentido unifilamentar pouco tem efeito na expressão do gene *mex-3*. Os embriões de vermes injetados com RNA bifilamentar de *mex-3* não levam nenhum RNA mensageiro detectável de *mex-3*. A ausência de RNA mensageiro de *mex-3* nesses embriões é o resultado de RNA de interferência induzido pelo RNA bifilamentar injetado. Os embriões dos vermes injetados com RNA anti-sentido unifilamentar de *mex-3* levavam algum RNA mensageiro de *mex-3*. Assim, o RNA anti-sentido unifilamentar de *mex-3* não é tão eficiente quanto o RNA bifilamentar de *mex-3* na indução de RNAi.

## Questões e Problemas

Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 21.1. Óperons são comuns em bactérias mas não em eucariontes. Sugira um motivo.
- 21.2. Usando exemplos, descreva as dimensões espaciais e temporais da regulação gênica eucariótica.
- 21.3. Que fatores influenciam a concentração e a longevidade das moléculas de mRNA?
- 21.4. Por que hormônios esteróides interagem com receptores dentro da célula, enquanto hormônios peptídicos interagem com receptores na superfície celular?
- 21.5. Como você pode usar os cromossomos politênicos de insetos dípteros para estudar a regulação da transcrição?
- 21.6. Como você pode distinguir um acentuador de um promotor?



- 21.7. Tropomiosinas são proteínas que medeiam a interação de actina e troponina, duas proteínas envolvidas na contração muscular. Em animais superiores, tropomiosinas existem como uma família de proteínas muito correlatas que compartilham algumas seqüências de aminoácidos mas diferem em outras. Explique como estas proteínas poderiam ser criadas a partir do transcrito de um único gene.
- 21.8. Em *Arabidopsis*, um gene de  $\beta$ -tubulina (*TUB1*) é expresso nas raízes e outro gene de  $\beta$ -tubulina (*TUB8*) é expresso nos tecidos vasculares. Sugira um mecanismo para explicar esta expressão histo específica.
- 21.9. Que técnicas podem ser usadas para mostrar que um gene de planta é transcrito quando a planta é iluminada com luz?
- 21.10. Um pseudogene especifica um RNA? Ele especifica um polipeptídeo funcional?
- 21.11. Usando técnicas de engenharia genética, um pesquisador construiu um gene de fusão contendo os elementos de resposta a choque provocado por calor de um gene *hsp70* de *Drosophila* e a região codificante de um gene de água-viva (*gfp*) para proteína fluorescente verde. Esse gene de fusão foi inserido nos cromossomos de drosófilas vivas por técnicas de transformação mediada por transposons (Cap. 18). Em que condições a proteína fluorescente verde será produzida nessas moscas geneticamente transformadas? Explique.
- 21.12. Suponha que o segmento do gene *hsp70* que foi usado para fazer a fusão *hsp70/gfp* no problema precedente tinha mutações em cada um de seus elementos de resposta ao choque provocado por calor. A proteína fluorescente verde codificada por este gene de fusão seria produzida em moscas geneticamente transformadas?
- 21.13. Por que faz sentido que a enzima fotossintética ribulose 1,5-bisfosfato-carboxilase (RBC) seja produzida especificamente quando plantas são expostas a luz?
- 21.14. Diferencie os motivos *zinc finger* e zíper de leucina encontrados em alguns fatores de transcrição eucarióticos.
- 21.15. O produto protéico do gene *sex-lethal* de *Drosophila* é necessário em fêmeas mas não em machos. Preveja o fenótipo de uma *Drosophila* XY que seja hemizigota para uma mutação nula no gene ligado ao X *Sex-lethal*.
- 21.16. As formas recompostas alternativamente de RNA do gene *doublesex* de *Drosophila* codificam proteínas que são necessárias para bloquear o desenvolvimento de um ou outro conjunto de características sexuais. A proteína que é feita em fêmeas bloqueia o desenvolvimento de características masculinas e a proteína que é feita em machos bloqueia o desenvolvimento de características femininas. Preveja o fenótipo dos animais XX e XY homozigotos para uma mutação nula no gene *doublesex*.
- 21.17. Em *Drosophila*, a expressão do gene *yellow* é necessária para a formação do pigmento escuro em muitos tecidos diferentes; sem esta expressão, um tecido apresenta-se de cor amarela. Nas asas, a expressão do gene *yellow* é controlada por um acentuador situado antecedente ao sítio de iniciação da transcrição do gene. Nas patas do tarso, a expressão é controlada por um acentuador situado dentro do único íntron do gene. Suponha que, por engenharia genética, o acentuador da asa é colocado dentro do íntron e o acentuador das patas é colocado antecedente ao sítio de início de transcrição. Uma mosca levando este gene *yellow* modificado em lugar do seu gene *yellow* natural teria asas e patas de pigmento mais escuro? Explique.
- 21.18. Descreva a evidência citológica de que o RNA é preferencialmente transcrito a partir de segmentos "frouxamente" organizados dos cromossomos.
- 21.19. Qual a natureza de cada uma das seguintes classes de enzimas? O que isto tem a ver com cromatina? (a) HAT, (b) HDAC, (c) HMT.
- 21.20. Em larvas de *Drosophila*, o único cromossomo X em machos parece difuso e borrado nas células politênicas da glândula salivar. Esta observação é compatível com a idéia de que os genes ligados ao X são hiperativados em machos de *Drosophila*?
- 21.21. Que modificações químicas da histona H4 estão associadas a genes transcricionalmente ativos?
- 21.22. O RNA bifilamentar derivado de um íntron seria capaz de induzir o RNA de interferência?
- 21.23. Que enzimas estão envolvidas na maturação de pequenos RNA de interferência e microRNA?
- 21.24. Que nucleotídeos são metilados no DNA de mamíferos?
- 21.25. O que é *imprinting*?
- 21.26. Suponha que fêmeas de camundongos homozigotas para o alelo *a* do gene *Igf2* são cruzadas com camundongos machos homozigotos para o alelo *b* deste gene. Quais destes dois alelos serão expressos na prole F<sub>1</sub>?
- 21.27. Cite um exemplo de amplificação gênica no qual cópias extracromossômicas do gene são criadas para fornecer mais moldes para transcrição.
- 21.28. Em tripanossomas, que mecanismos resultam na expressão de uma glicoproteína variante de superfície diferente?
- 21.29. Um pesquisador supôs que em camundongos o gene *A* é ativamente transcrito em células hepáticas, enquanto o gene *B* é ativamente transcrito em células do cérebro. Descreva os procedimentos que permitiriam ao pesquisador testar esta hipótese.
- 21.30. Suponha que a hipótese mencionada na questão anterior é correta e que o gene *A* é ativamente transcrito em células do



figado enquanto o gene *B* é ativamente transcrito em células do cérebro. O pesquisador agora extrai quantidades equivalentes de cromatina de tecidos do fígado e do cérebro e trata esses extratos separadamente com DNase I por um período de tempo limitado. Se o DNA que sobra após o tratamento é então fracionado por eletroforese em gel, transferido para uma membrana por transferência de Southern e hibridizado com uma sonda radioativamente marcada específica do gene

*A*, que amostra (fígado ou cérebro) se esperará apresentar o maior sinal no auto-radiograma? Explique sua resposta.

**21.31.** Por que mutações nulas no gene *mle* em *Drosophila* não têm efeito em fêmeas?

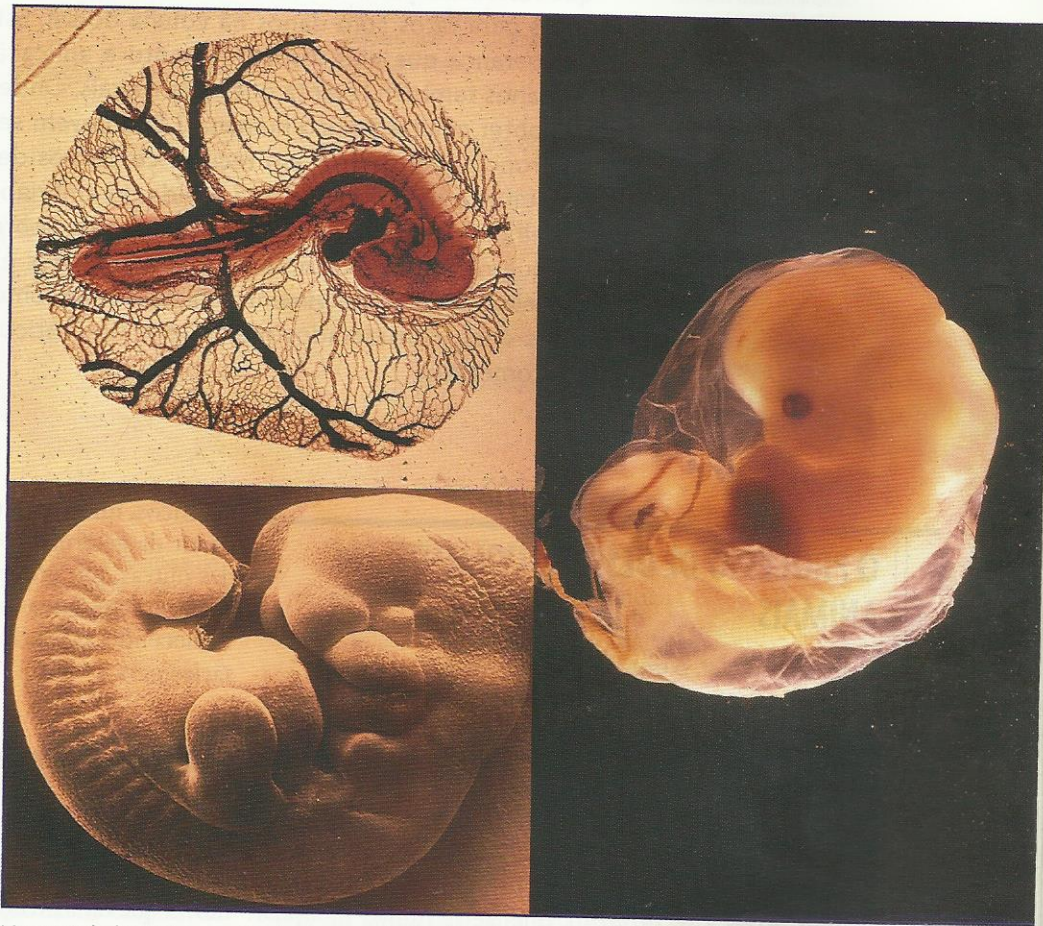
**21.32.** Quais são os processos envolvidos em inativação do cromossomo X em mamíferos?



# O Controle Genético do Desenvolvimento Animal

## Tópicos

- O Processo de Desenvolvimento em Animais
- Análise Genética do Desenvolvimento em Organismos Modelos
- Análise Genética das Vias de Desenvolvimento
- Atividade Gênica Materna no Desenvolvimento
- Atividade Gênica Zigótica no Desenvolvimento
- Análise Genética do Desenvolvimento em Vertebrados



No sentido horário a partir de cima: embrião de galinha 56 horas após fertilização; embrião humano oito semanas após fertilização; embrião de camundongo 10 dias após fertilização.

## Terapia Com Células Tronco: Um Admirável Mundo Novo?

As células tronco estão nos noticiários. Cientistas estão discutindo seus possíveis usos, e todos os tipos de pessoas – políticos, líderes religiosos, jornalistas e vítimas de doenças tais como doença de Parkinson, diabete e artrite, e mesmo celebridades de Hollywood – juntaram-se ao debate. Duas



características definem uma célula tronco. Primeiro, ela é capaz de aparentemente reproduzir-se sem limite, divisão após divisão, produzindo descendentes iguais a ela. Segundo, ela é capaz de gerar uma prole comprometida com diferenciação em tipos especiais de células, como fibras musculares, linfócitos, neurônios ou células musculares. Devido a estas duas características, as células tronco podem ter o potencial de múltiplos usos terapêuticos: regenerar tecidos desgastados, substituir órgãos perdidos ou partes do corpo, corrigir danos na coluna dorsal ou no cérebro, aliviar déficits bioquímicos. Embora para seres humanos estas possibilidades ainda sejam ficção científica, existem evidências experimentais de que a terapia com células tronco pode funcionar. Células tronco derivadas de embriões de camundongo podem formar células de apoio neuronal que compensam um defeito bioquímico quando são transplantadas para cérebros de camundongo. Elas também podem formar neurônios que restauram parcialmente o funcionamento da coluna dorsal quando são transplantadas para ratos com danos na coluna dorsal. Estes resultados são promissores, mas a esperança que eles inspiram quanto à aplicação da terapia de células tronco em humanos também levanta questões morais. Onde pesquisadores e médicos conseguirão obter células tronco humanas? Elas devem ser derivadas de embriões ou as células tronco que residem em tecidos corpóreos de adultos devem ser adequadas para pesquisas e fins terapêuticos? A vida do embrião deve ser sacrificada para prolongar e melhorar a vida adulta? O intenso debate sobre terapia de células tronco demonstra que os avanços tecnológicos baseados nos estudos de diferenciação celular e o processo de desenvolvimento animal estão começando a criar problemas éticos que antes eram considerados meramente hipotéticos. O Admirável Mundo Novo pode ter começado.

## O PROCESSO DE DESENVOLVIMENTO EM ANIMAIS

A divisão celular, o movimento celular e a especialização celular têm papéis na formação de um animal multicelular a partir de um zigoto.

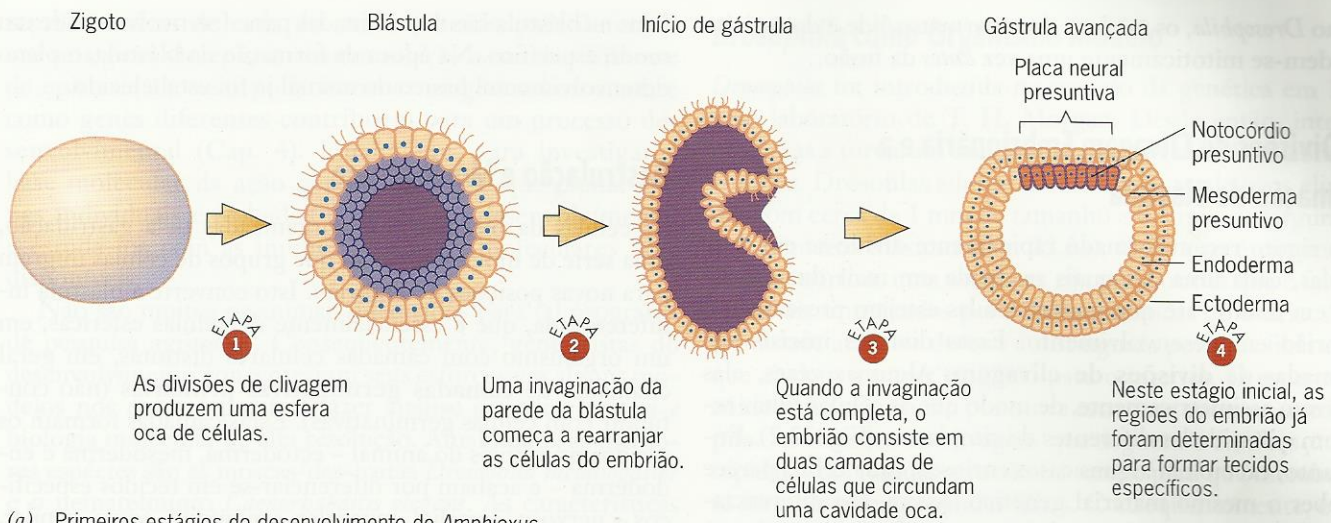
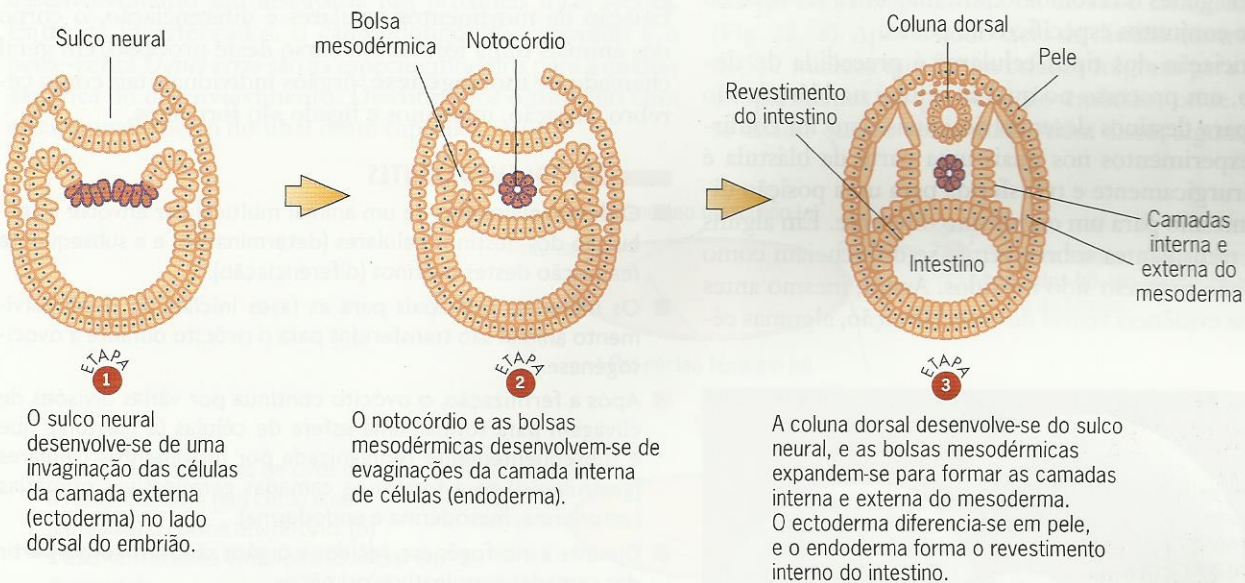
O desenvolvimento animal na verdade começa antes da fertilização do ovócito. Substâncias críticas são sintetizadas pela fêmea e seqüestradas no ovócito para nutrir e controlar o desenvolvimento do zigoto em um organismo multicelular. Após a fertilização, uma série rápida de divisões mitóticas aumenta o número de células no animal em desenvolvimento, geralmente sem muito crescimento no tamanho geral (Fig. 22.1). A massa resultante de células, a **blástula**, é relativamente indistinta, mas, mesmo neste estágio, algumas de suas células foram marcadas para formar tecidos específicos no corpo do animal. Tais células foram destinadas, ou determinadas, a se desenvolver de um certo modo. À medida que o desenvolvimento continua, a blástula é reorganizada por uma série de migrações celulares, coletivamente chamadas de **gastrulação**, para produzir um embrião com três camadas celulares distintas – **ectoderma** (do lado externo), **mesoderma** (no meio) e **endoderma** (no interior). Estas três camadas dão origem a tecidos es-

pecíficos, que são então estruturados para formar órgãos. A partir da gastrulação, as células tornam-se diferentes umas das outras, um processo chamado de **diferenciação**. Esta é uma característica importante do desenvolvimento animal, fornecendo as várias funções especializadas que são necessárias para a vida.

## Ovocitogênese e Fertilização

Todo o potencial do desenvolvimento de um animal está contido em um zigoto. Quando espermatozóide e ovócito se unem, eles juntam os conjuntos paternos e maternos de genes, o que dá instruções detalhadas para o desenvolvimento. Entretanto, o desenvolvimento também é profundamente afetado pelos materiais citoplasmáticos presentes no ovócito. Alguns destes materiais, tais como o vitelo, são nutritivos e servem de alimento para o animal em desenvolvimento. Outros materiais são determinantes em ajudar a estabelecer o próprio padrão básico de desenvolvimento. Tanto os componentes nutritivos quanto os determinantes do citoplasma do ovo são fabricados durante a ovocitogênese, o processo de formação do ovócito. Isto envolve uma série complexa de eventos, incluindo as duas divisões meióticas, que reduzem o número de cromossomos de diplóide para haplóide. Curiosamente, a segunda divisão meiótica às vezes não é completada senão após a fertilização.



(a) Primeiros estágios do desenvolvimento de *Amphioxus*.(b) Morfogênese em *Amphioxus*.**Fig. 22.1** ■ Desenvolvimento embrionário do animal marinho *Amphioxus*, um protocordado. (a) Primeiros estágios do desenvolvimento embrionário. (b) Morfogênese.

Em fêmeas, as divisões meióticas produzem células de tamanhos desiguais (Cap. 2). A primeira divisão separa uma célula grande, o ovócito secundário, de uma célula pequena, o primeiro glóbulo polar. O ovócito secundário desenvolve-se no óvulo, enquanto o glóbulo polar degenera. Esta divisão desigual de células é claramente um modo de estocar materiais no citoplasma do zigoto. A segunda divisão meiótica também é desigual, pois a maior parte do citoplasma é reservada para o óvulo; apenas pequena parte formará o segundo glóbulo polar.

As células vizinhas também nutrem o citoplasma do óvulo. Em muitos organismos, estas células produzem materiais que são transportados para o ovócito, tendo uma função nutritiva ou de apoio. As células vizinhas também podem

fornecer materiais para envolver o ovócito. Em aves, por exemplo, uma envoltório duro rico em cálcio é formado por secreções no oviduto, uma parte da via reprodutiva feminina pela qual o ovo passa em seu caminho para fora do corpo.

A fertilização ocorre quando o ovócito é penetrado por um espermatozoide. Este é um processo complexo, envolvendo interações de proteínas diferentes nas superfícies dos dois gametas. Quando o espermatozoide faz contato com o ovócito, ele ativa uma resposta no ovócito que faz com que o citoplasma perto da superfície do ovócito envolva o espermatozoide e o internalize. Após o espermatozoide ter entrado no ovócito, este completa a meiose (se já não o tiver feito) e os núcleos do espermatozoide e do ovócito fundem-se para formar o núcleo do zigoto. Em alguns organismos,



como *Drosophila*, os núcleos do espermatozóide e do ovócito dividem-se mitoticamente uma vez *antes* da fusão.

### As Divisões de Clivagem Embrionária e a Formação de Blástula

Um zigoto recém-formado rapidamente divide-se em duas células, cada uma das quais se divide em mais duas, e assim em diante, até que muitas células estejam presentes no embrião em desenvolvimento. Essas divisões iniciais são chamadas de **divisões de clivagem**. Algumas vezes, elas ocorrem assimetricamente, de modo que as células filhas recebem quantidades diferentes de citoplasma (Fig. 22.2). Entretanto, exceto em alguns casos curiosos, cada célula parece receber o mesmo material genético. Assim, sua diferenciação subsequente em tipos celulares específicos não se deve à segregação de genes ou cromossomos, mas sim a ativação ou repressão de conjuntos específicos de genes.

A diferenciação dos tipos celulares é precedida de **determinação**, um processo no qual as células na blástula são orientadas para destinos desenvolvimentais. Isto foi confirmado por experimentos nos quais uma parte da blástula é removida cirurgicamente e transferida para uma posição diferente, ou mesmo para um organismo diferente. Em alguns casos, esses transplantes sobrevivem e se diferenciam como o fariam se não tivessem sido movidos. Assim, mesmo antes de haver uma evidência visível de especialização, algumas cé-

lulas na blástula são determinadas para desenvolver-se de um modo específico. Na época da formação da blástula, o plano desenvolvimental básico do animal já foi estabelecido.

### Gastrulação e Morfogênese

As células da blástula são reorganizadas pela gastrulação, uma série de movimentos em que grupos de células migram para novas posições no embrião. Isto converte a blástula indiferenciada, que é essencialmente de células esféricas, em um organismo com camadas celulares distintas, em geral chamadas de **camadas germinativas primárias** (não confundir com células germinativas). Estas camadas formam os tecidos primitivos do animal – ectoderma, mesoderma e endoderma – e acabam por diferenciar-se em tecidos específicos – nervoso, epitelial, muscular, sangüíneo, ósseo, capilar, glandular, visceral, e assim em diante. Mediante uma combinação de movimentos celulares e diferenciação, o corpo dos animais toma forma. No curso deste processo, em geral chamado de **morfogênese**, órgãos individuais tais como cérebro, coração, intestinos e fígado são formados.

#### PONTOS IMPORTANTES

- O desenvolvimento de um animal multicelular envolve a atribuição dos destinos celulares (determinação) e a subsequente realização destes destinos (diferenciação).
- Os materiais principais para as fases iniciais do desenvolvimento animal são transferidos para o ovócito durante a ovotogênese.
- Após a fertilização, o ovócito continua por várias divisões de clivagem para formar uma esfera de células (a blástula), que é subsequentemente reorganizada por movimentos celulares (gastrulação) para formar as camadas germinativas primárias (ectoderma, mesoderma e endoderma).
- Durante a morfogênese, tecidos e órgãos são formados a partir das camadas germinativas primárias.



**Fig. 22.2** ■ Micrografia eletrônica de varredura de um embrião do caracol marinho *Ilyanassa obsoleta* no estágio de oito células. A pequena célula no centro é um glóbulo polar.

### ANÁLISE GENÉTICA DO DESENVOLVIMENTO EM ORGANISMOS MODELOS

*Drosophila* e *C. elegans* foram os primeiros organismos modelos para o estudo do desenvolvimento animal.

A biologia desenvolvimental moderna depende intensamente de análise genética, que deve começar com a identificação de genes relevantes. Isto é feito coletando-se mutações com fenótipos diagnósticos; por exemplo, se desejamos estudar o desenvolvimento das asas de um inseto, começaríamos coletando mutações que alteram ou evitam o desenvolvimento de asas. Então mapearíamos estas mutações e as testaríamos quanto a alelismo umas com as outras para definir e posicionar os *loci* genéticos relevantes. Uma vez identificados esses *loci*, combinaríamos mutações



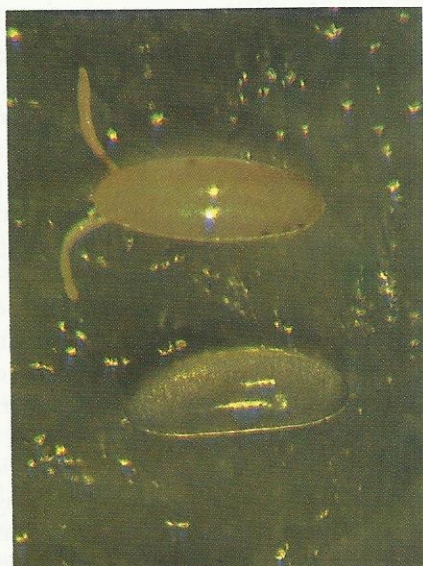
representativas de cada de modo pareado para determinar se alguns são epistáticos em relação a outros. Tal teste de epistasia às vezes pode fornecer *insights* valiosos sobre como genes diferentes contribuem para um processo desenvolvimental (Cap. 4). Finalmente, para investigar a base molecular da ação gênica, teríamos que clonar genes individuais envolvidos no processo desenvolvimental e estudá-los com as inúmeras técnicas moleculares hoje disponíveis.

Não são muitos os animais adequados para tal programa de pesquisa agressivo. Conseqüentemente, geneticistas do desenvolvimento concentraram seus esforços em alguns modelos nos quais é possível fazer análise genética intensa e biologia molecular de alta resolução. Atualmente, as melhores espécies são as moscas-das-frutas *Drosophila melanogaster* e o nematelminto *Caenorhabditis elegans*. As características que tornam estes dois invertebrados ideais para o estudo do desenvolvimento são discutidas nas próximas duas seções. Entre os invertebrados, o camundongo *Mus musculus* e o peixe-zebra *Danio rerio* são as espécies modelos para a análise genética do desenvolvimento. Discutiremos o trabalho com estes animais perto do final deste capítulo.

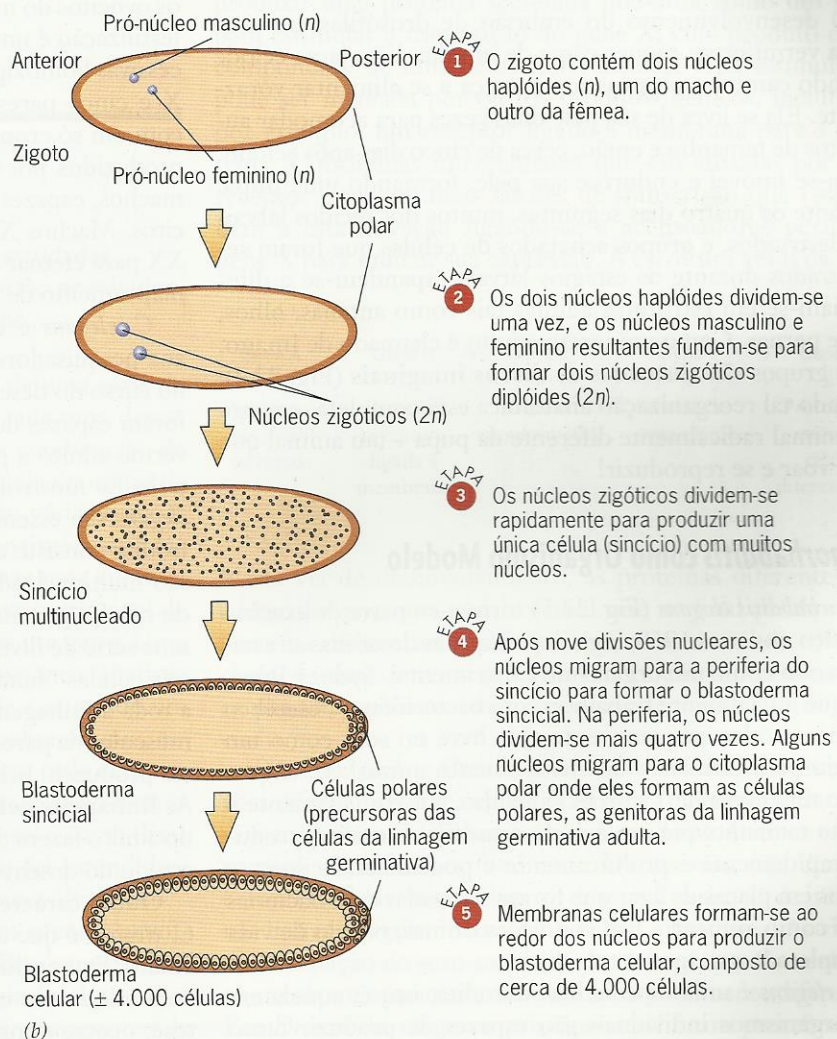
## *Drosophila* como Organismo Modelo

*Drosophila* foi introduzida no mundo da genética em 1909 pelo laboratório de T. H. Morgan. Desde então, intensas pesquisas a tornaram um dos eucariontes mais bem compreendidos. *Drosófilas* adultas desenvolvem-se de ovos elipsóides com cerca de 1 mm de tamanho e 0,5 mm de largura em seu diâmetro máximo (Fig. 22.3a). Cada ovo é circundado por um **córior**, uma estrutura rija tipo concha que é feita de materiais sintetizados por células somáticas no ovário. A extremidade anterior do ovo é distinguida por dois filamentos que ajudam a trazer oxigênio para o interior. O espermatozóide entra no ovócito por outra estrutura anterior, a **micrópila**. As divisões de clivagem que seguem a fertilização são rápidas – tão rápidas que não há tempo para que membranas se formem entre as células filhas. Conseqüentemente, o embrião inicial de *Drosophila* é de fato uma única célula com muitos núcleos idênticos. Tal célula é chamada de **sincício** (Fig. 22.3b). Após o nono ciclo de divisão dentro do sincício, os 512 núcleos que foram criados migram para a membrana citoplasmática na periferia do embrião, onde continuam a se dividir quatro vezes mais. Além disso, alguns dos núcleos

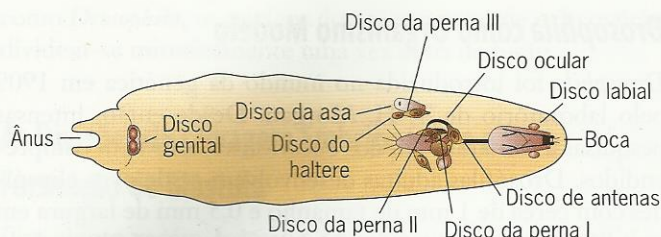
**Fig. 22.3** ■ (a) Foto de ovos de *Drosophila*, com (em cima) e sem (embaixo) o córior envoltório. (b) Desenvolvimento embrionário inicial em *Drosophila*.



(a)







**Fig. 22.4** ■ A localização de discos imaginais em larvas de *Drosophila*.

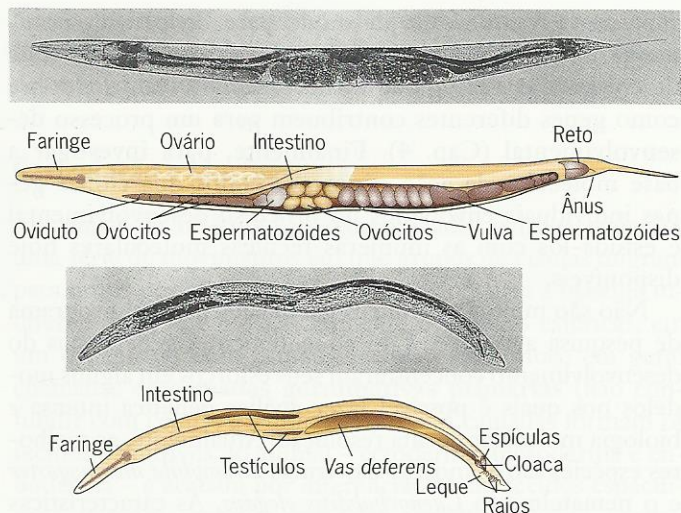
migram para o pólo posterior do embrião. No 13º ciclo de divisão, todos os núcleos do sincício tornam-se separados uns dos outros por membranas celulares, criando uma camada única de células na superfície do embrião. Esta camada única, chamada de **blastoderma celular**, é equivalente à blástula de outros organismos. Ela dará origem a todos os tecidos somáticos do animal. A celularização dos núcleos no pólo posterior cria as **células polares**, que originam a linhagem germinativa adulta. Assim, neste estágio inicial do desenvolvimento, as linhagens somática e germinativa do futuro adulto já se separaram.

O desenvolvimento do embrião de drosófilas em uma **larva** vermiforme requer cerca de um dia. Esta larva eclode abrindo caminho pela casca e começa a se alimentar vorazmente. Ela se livra de sua pele duas vezes para acomodar aumentos de tamanho e então, cerca de cinco dias após eclodir, torna-se imóvel e endurece sua pele, formando uma **pupa**. Durante os quatro dias seguintes, muitos dos tecidos larvais são destruídos, e grupos achatados de células que foram seqüestrados durante os estágios larvais expandem-se e diferenciam-se em estruturas adultas tais como antenas, olhos, asas e pernas. Como um inseto adulto é chamado de **imago**, esses grupos são chamados de **discos imaginais** (Fig. 22.4). Quando tal reorganização anatômica está completa, emerge um animal radicalmente diferente da pupa – um animal que pode voar e se reproduzir!

### **Caenorhabditis como Organismo Modelo**

*Caenorhabditis elegans* (Fig. 22.5) tornou-se parte do cenário genético nos anos 1960, logo depois que as drosófilas se estabeleceram como um organismo experimental. Sydney Brenner, que inicialmente trabalhou com bacteriófagos, escolheu este nematódeo pequeno e de vida livre no solo como um modelo para estudar o desenvolvimento animal. *C. elegans* adulto tem cerca de 1 mm de tamanho, aproximadamente o mesmo tamanho que um ovo de drosófila. Eles se reproduzem rapidamente e prolificamente e podem ser facilmente criados em placas de ágar que foram semeadas com bactérias *E. coli* como alimento. Em condições ótimas, o ciclo de vida é completado em apenas três dias.

*C. elegans* é uma espécie hermafrodita, isto é, aquela em que organismos individuais são capazes de produzir tanto



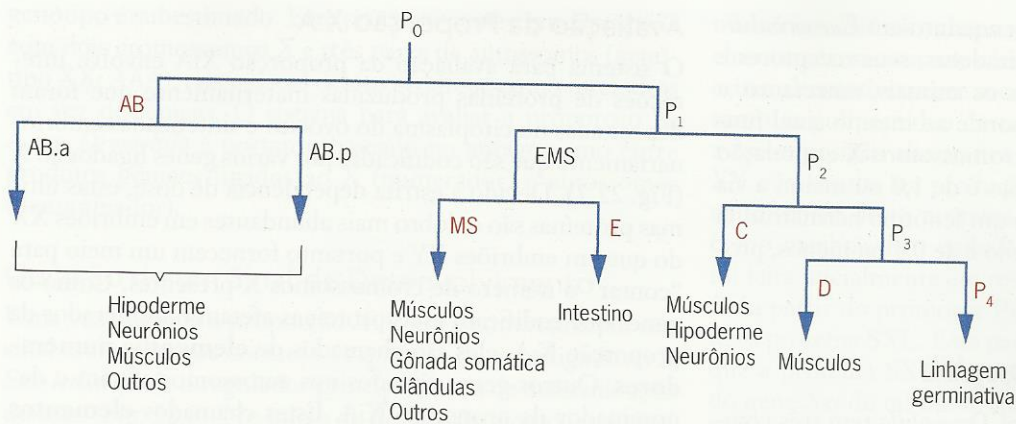
**Fig. 22.5** ■ Fotos e diagramas mostrando os fenótipos de um *Caenorhabditis* hermafrodita (em cima) e masculino (embaixo).

espermatozoides quanto ovócitos. Como não há sistema de auto-incompatibilidade, os espermatozoides podem fertilizar os ovócitos do mesmo animal. Para o pesquisador, essa autofertilização é um modo conveniente de tornar mutações recessivas homozigotas. Hermafroditas têm dois cromossomos X e cinco pares de autossomos. Ocasionalmente, animais com um só cromossomo X e cinco pares de autossomos são produzidos por não-disjunção meiótica. Esses animais são machos, capazes de produzir espermatozoides mas não ovócitos. Machos XO podem ser cruzados com hermafroditas XX para efetuar procedimentos genéticos padrões tais como mapeamento de recombinação e testes de complementação.

*C. elegans* é um animal transparente, que proporciona aos pesquisadores a oportunidade de observar cada célula no curso do desenvolvimento. John Sulston e colaboradores foram capazes de rastrear a linhagem de todas as células no verme adulto a partir do zigoto unicelular (Fig. 22.6). Este trabalho mostrou que os eventos do desenvolvimento de *C. elegans* são essencialmente invariantes. Cada adulto hermafrodita consiste em 959 núcleos somáticos (algumas células são multinucleadas) bem como um número indeterminado de células germinativas. Um zigoto de *C. elegans* passa por uma série de divisões assimétricas de clivagem para produzir seis células "fundadoras". Uma delas acaba por dar origem a toda a linhagem germinativa, uma ao intestino e uma aos músculos da parede corpórea. As outras três células fundadoras produzem principalmente células nervosas e musculares. As **linhagens celulares invariantes** que formam os tecidos do adulto fazem de *C. elegans* um excelente organismo para o estudo do desenvolvimento.

Outra característica do desenvolvimento evidente em *C. elegans* é que algumas células não sobrevivem ao estágio adulto. Estas células são programadas para morrer. Tal fenômeno de morte celular, que os cientistas chamam de apoptose, ocorre durante o desenvolvimento de muitos animais.





**Fig. 22.6** ■ Uma parte da linhagem celular de *C. elegans* hermafrodita.  $P_0$ , o zigoto, divide-se para produzir duas células indicadas como AB e  $P_1$ . Cada uma delas então se divide para produzir dois pares de células, um par indicado como AB.a e AB.p e o outro indicado como EMS e  $P_2$ . Divisões posteriores destas células e suas descendentes produzem todas as células do adulto hermafrodita. As seis células fundadoras são indicadas em vermelho.

No Cap. 24, veremos como isto também tem um papel em evitar a formação de cânceres.

### PONTOS IMPORTANTES

- Em drosófilas, a sequência de desenvolvimento é zigoto, embrião, larva, pupa, adulto.
- As linhagens de todas as células em um *C. elegans* adulto podem ser rastreadas até o zigoto unicelular.

## ANÁLISE GENÉTICA DAS VIAS DE DESENVOLVIMENTO

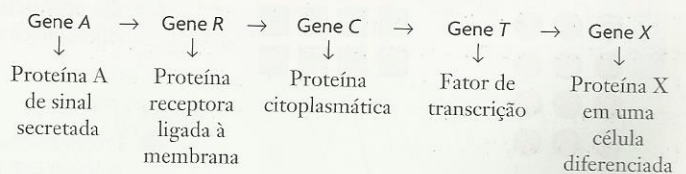
As vias de desenvolvimento podem ser estudadas identificando-se genes cujos produtos estão envolvidos na diferenciação de fenótipos específicos.

A noção de que os fenótipos resultam de uma série de etapas em uma via e que genes controlam cada uma destas etapas surgiu nos anos 1930 e 1940 de uma combinação de investigações genéticas e bioquímicas. Naquela época, já tinha sido estabelecido que as células efetuam várias reações bioquímicas, cada uma catalisada por uma enzima, e que várias reações podem ser associadas para formar uma via (veja Cap. 14). A genética proporcionou a percepção principal de que cada enzima em uma via é codificada por um gene. Uma mutação em um desses genes poderia, portanto, inativar uma enzima essencial, bloqueando toda a via e causando um fenótipo mutante. Os geneticistas logo perceberam que, estudando tais mutações, eles podiam identificar cada etapa na via. Além disso, analisando pares de mutações, eles podiam às vezes dispor as etapas em uma ordem temporal correta. Essa dissecação genética de vias bioquímicas logo foi estendida a outros processos biológicos, incluindo o desenvolvimento.

Uma via de desenvolvimento consiste em eventos envolvidos na diferenciação de tecidos e órgãos. Produtos gênicos diferentes participam desses eventos, incluindo alguns que são moléculas sinalizadoras, alguns que são receptores de

sinal, outros que são transdutores de sinal e ainda outros que são fatores de transcrição. Outros tipos de proteínas regulatórias também podem estar envolvidos. Ao final, uma via gera os componentes que formam as estruturas de tecidos e órgãos particulares. Ela cria, portanto, um fenótipo.

Os componentes de uma via são ordenados causalmente pelo modo em que causam o fenótipo. O gene *A* pode produzir uma proteína secretada que atua como um sinal para estimular a transcrição do gene *X*, cujo produto é um componente de uma célula diferenciada. Essa estimulação pode ser mediada por outros produtos gênicos, incluindo, por exemplo, um receptor ligado a membrana para a proteína *A*, proteínas intracelulares que são ativadas por este receptor e um ou mais fatores de transcrição que respondem a esta ativação ligando-se a acentuadores perto do gene *X* para induzir sua expressão. A estrutura geral da via é



Nesta via de desenvolvimento, as proteínas diferentes são ordenadas por uma cadeia causal que ao final produz uma proteína característica de uma determinada célula diferenciada. As setas entre os genes na via indicam a ordem na qual os produtos gênicos atuam para causar um estado diferenciado.

Em *Drosophila* e *Caenorhabditis*, as vias desenvolvimentais para diferenciação sexual estão entre as mais estudadas. Tais vias controlam a diferenciação dos tecidos somáticos em estruturas masculinas e femininas. Outras vias, não tão bem compreendidas, controlam o desenvolvimento de tecidos nas linhagens germinativas masculina e feminina. As vias de determinação do sexo somático de *Drosophila* e *Caenorhabditis* envolvem mecanismos moleculares totalmente diferentes. Em *Drosophila*, os genes principais codificam proteínas que



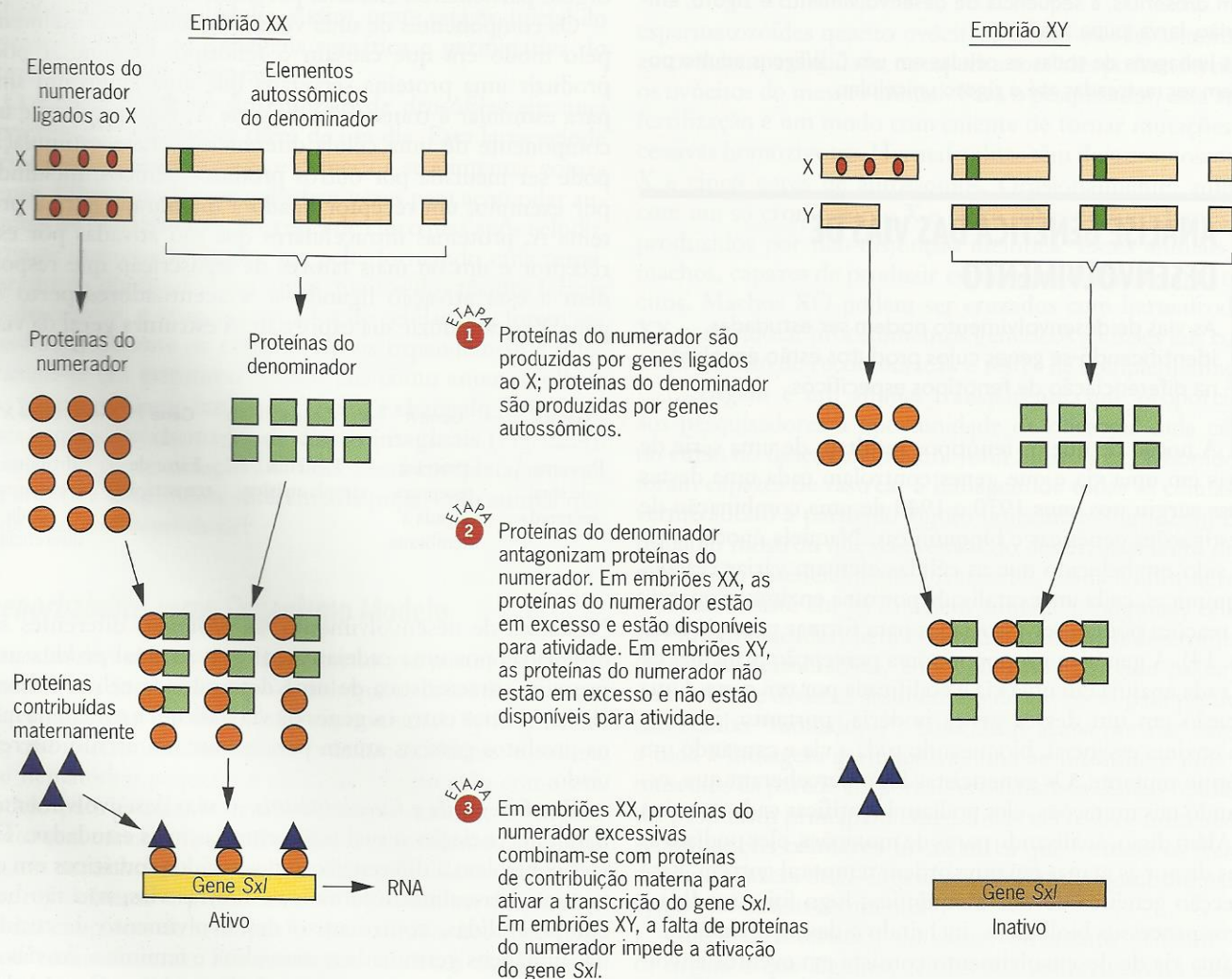
regulam recomposição do RNA, enquanto em *Caenorhabditis* elas codificam moléculas sinalizadoras, seus receptores e fatores de transcrição. Em ambos os animais, entretanto, a via de determinação do sexo responde ao mesmo sinal fundamental, que é a proporção de cromossomos X em relação a autossomos. Quando a proporção é de 1,0 ou maior, a via produz um fenótipo feminino (ou um fenótipo hermafrodita em *C. elegans*), ao passo que, quando é de 0,5 ou menos, produz um fenótipo masculino.

### Determinação do Sexo em *Drosophila*

A via de determinação do sexo em *Drosophila* tem três componentes: (1) um sistema para avaliar a proporção X:A bem no início do embrião; (2) um sistema que converte esta proporção em um sinal desenvolvimental; e (3) um sistema que responde a este sinal produzindo estruturas masculinas ou femininas.

### Avaliação da Proporção X:A

O sistema para avaliação da proporção X:A envolve interações de proteínas produzidas maternamente que foram depositadas no citoplasma do ovócito e sintetizadas embrionicamente que são codificadas por vários genes ligados ao X (Fig. 22.7). Devido à estrita dependência de dose, estas últimas proteínas são o dobro mais abundantes em embriões XX do que em embriões XY e portanto fornecem um meio para “contar” o número de cromossomos X presentes. Como os genes que codificam estas proteínas afetam o numerador da proporção X:A, eles são chamados de **elementos numeradores**. Outros genes situados nos autossomos afetam o denominador da proporção X:A. Estes chamados **elementos denominadores** codificam proteínas que antagonizam os produtos dos elementos numeradores. Conseqüentemente, à medida que a dosagem dos elementos do denominador aumenta, a dosagem “percebida” de elementos do numerador é diminuída, e o número de cromossomos X presentes no



**Fig. 22.7** ■ Avaliação da proporção X:A por elementos do numerador e do denominador em *Drosophila*. A proporção de cromossomos X em relação a autossomos é avaliada por interações dos produtos protéicos destes genes.



genótipo é subestimado. Este processo ocorre em *Drosophila* com dois cromossomos X e três pares de autossomos (genótipo XX; AAA); tais moscas desenvolvem-se como intersexos em vez de fêmeas. O sistema para avaliar a proporção X:A em *Drosophila* é portanto baseado no antagonismo entre produtos gênicos ligados ao X (numerador) e autossomos (denominador).

### Geração de um Sinal de Desenvolvimento

Uma vez avaliada a proporção X:A, ela é convertida em um sinal molecular que controla a expressão do gene ligado ao X *Sex-lethal* (*Sxl*), o regulador principal da via de determinação do sexo (Fig. 22.8). No início do desenvolvimento, este sinal ativa a transcrição do gene *Sxl* a partir de  $P_E$ , o promotor “inicial” do gene, mas só em embriões XX. Os transcritos iniciais deste promotor são processados e traduzidos para produzir proteína Sex-lethal funcional, denominada SXL. Após apenas alguns ciclos de divisão, a transcrição a partir do promotor  $P_E$  é substituída pela transcrição a partir de outro promotor,  $P_M$ , o chamado promotor de manutenção do gene *Sxl*. Curiosamente, a transcrição a partir do promotor  $P_M$  também é iniciada em embriões XY. Entretanto, os transcritos a partir de  $P_M$  são processados corretamente apenas se proteína SXL estiver presente. Conseqüentemente, em embriões XY, em que esta proteína ainda não foi feita, os

transcritos *Sxl* são alternativamente recompostos para incluir um éxon com um códon de fim, e, quando estes transcritos alternativamente recompostos são traduzidos, eles geram um polipeptídeo curto sem função reguladora. Assim, a recomposição alternativa dos transcritos *Sxl* em embriões XY não resulta na produção da proteína SXL funcional, e na ausência desta proteína esses embriões desenvolvem-se como machos. Em embriões XX, em que a proteína SXL foi feita inicialmente em resposta ao sinal X:A, os transcritos *Sxl* a partir do promotor  $P_M$  são recompostos para produzir mais proteína SXL. Este padrão de expressão é mantido porque a proteína SXL regula a recomposição dos transcritos do gene *Sxl* de tal modo que eles codificam toda a proteína SXL. Em embriões XX, esta proteína é portanto um regulador positivo de sua própria síntese – um curioso mecanismo de *feedback* que mantém a expressão da proteína SXL em embriões XX e impede sua expressão em embriões XY.

A proteína SXL também regula a recomposição de transcritos de outro gene na via de determinação do sexo, *transformer* (*tra*) (Fig. 22.9). Esses transcritos podem ser processados de dois modos diferentes. Em machos cromossômicos, em que a proteína SXL está ausente, o aparelho de recomposição sempre deixa um códon de fim no segundo éxon do RNA *tra*. Assim, quando o RNA *tra* recomposto é traduzido, ele gera um polipeptídeo truncado (e não funcional). Em fêmeas, em que a proteína SXL está presente, este códon de

#### ETAPA 1 Transcrição:

Em embriões XX, um sinal molecular baseado na proporção X:A inicia a transcrição do gene *Sxl* a partir do promotor  $P_E$ .

Depois, a transcrição é iniciada no promotor  $P_M$  nos embriões XX e XY.

#### ETAPA 2 Recomposição:

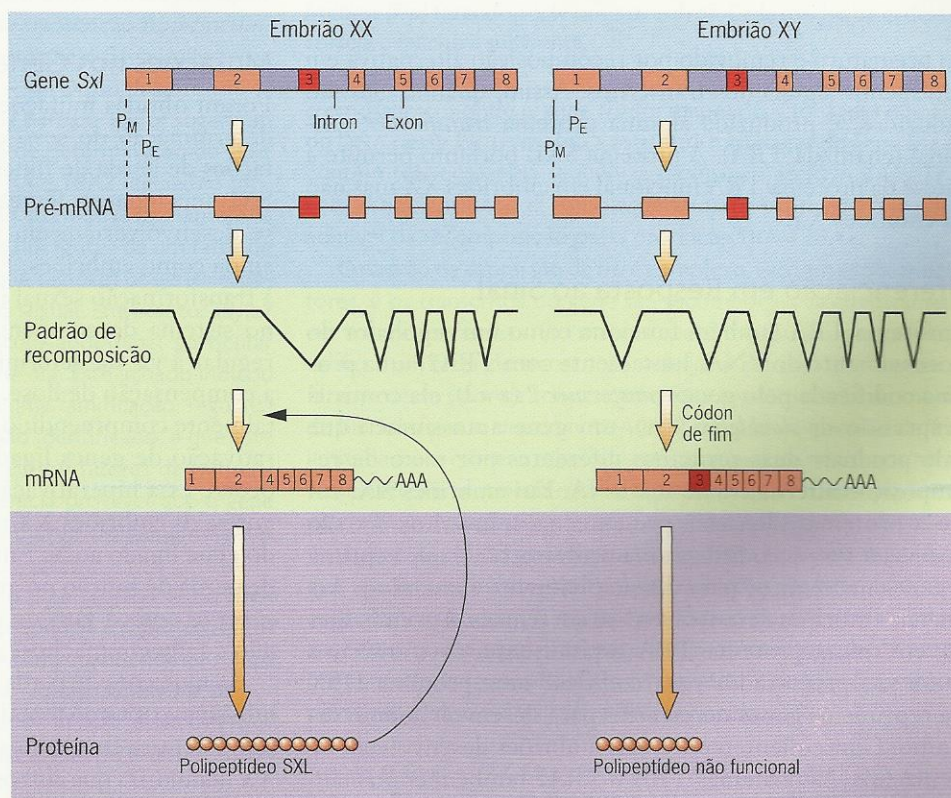
Em embriões XX, os transcritos *Sxl* são recompostos para conter todos os éxons exceto o éxon 3.

Em embriões XY, os transcritos de *Sxl* são recompostos para conter todos os éxons, incluindo o éxon 3.

#### ETAPA 3 Tradução:

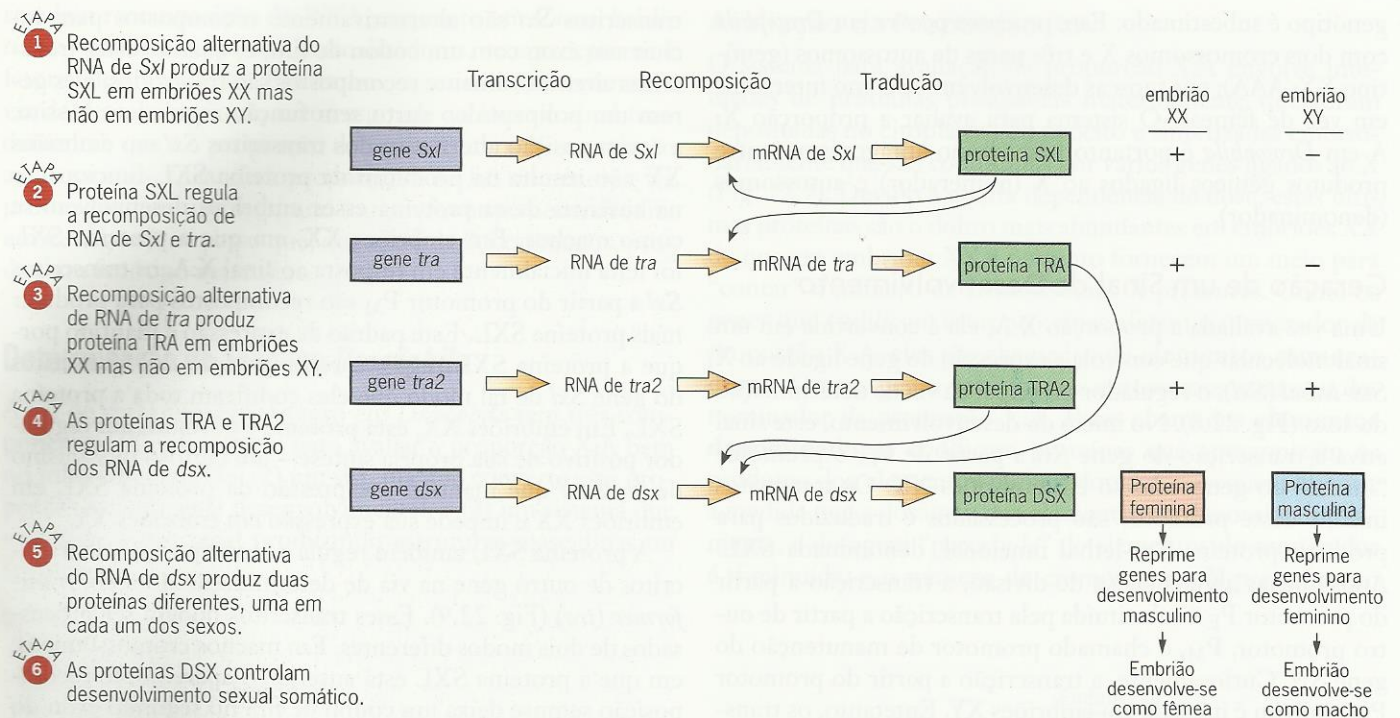
Em embriões XX, o mRNA de *Sxl* é traduzido em um polipeptídeo (SXL) que regula a recomposição, incluindo a recomposição de transcritos de *Sxl*.

Em embriões XY, um códon de fim no éxon 3 impede que o mRNA de *Sxl* seja traduzido em um polipeptídeo funcional.



**Fig. 22.8** ■ Expressão específica do sexo do gene *Sex-lethal* (*Sxl*) em *Drosophila*. Embora este gene seja transcrito tanto em embriões XX quanto em XY, a recomposição alternativa de seu RNA limita a síntese da proteína SXL a embriões XX, que se desenvolvem como fêmeas. A ausência da proteína SXL em embriões XY faz com que eles se desenvolvam como machos.





**Fig. 22.9** ■ Regulação da determinação do sexo em *Drosophila* pelo gene *Sex-lethal* (*Sxl*). O gene *Sxl* regula a expressão do gene *transformer* (*tra*), que, por sua vez, regula a expressão do gene *doublesex* (*dsx*); o gene *transformer2* (*tra2*) também participa na regulação de *dsx*. Os sinais + e - indicam a presença ou ausência das várias proteínas.

fim prematuro é removido por recomposição alternativa em pelo menos alguns dos transcritos. Assim, quando eles são traduzidos, é produzida alguma proteína *transformer* funcional (chamada TRA). A proteína SXL portanto permite a síntese da proteína TRA funcional em embriões XX mas não em embriões XY.

### Diferenciação em Resposta ao Sinal

A proteína TRA também funciona como um regulador do processamento do RNA. Juntamente com TRA2, uma proteína codificada pelo gene *transformer 2* (*tra2*), ela controla a expressão de *doublesex* (*dsx*), um gene autossômico que pode produzir duas proteínas diferentes por meio de recomposição alternativa de seu RNA. Em embriões XX, em que a proteína TRA está presente, os transcritos *dsx* são processados para codificar uma proteína DSX que reprime os genes necessários para desenvolvimento masculino. Assim, tais embriões desenvolvem-se em fêmeas. Em embriões XY, em que a proteína TRA está ausente, os transcritos de *dsx* são processados para codificar uma proteína DSX que reprime os genes necessários para desenvolvimento feminino. Conseqüentemente, tais embriões desenvolvem-se em machos. As proteínas TRA e TRA2 também controlam o processamento de transcritos de outro gene autossômico chamado *fruitless* (*fru*). Enfoque Técnico: *Fruitless* descreve como foi elucidado o papel deste gene na diferenciação sexual.

### Mutações nos Genes de Determinação Sexual

Foram obtidas mutações em cada um dos genes da via de determinação do sexo em *Drosophila* (Quadro 22.1). Mutações de perda de função em *Sxl* impedem que a proteína SXL seja feita em fêmeas. Mutantes homozigotos portanto se desenvolveriam em machos. Entretanto, eles morrem ainda como embriões. Essa morte embrionária não se deve à transformação sexual incipiente, mas sim a uma anomalia no sistema de compensação de dose (Cap. 21). Além de regular a via de determinação do sexo, *Sxl* também regula a compensação de dose. Embora o mecanismo não seja totalmente compreendido, ele aparentemente impede a hiperativação de genes ligados ao X em animais XX. Quando ocorre essa hiperativação, como em mutantes *Sxl* homozigotos, os embriões XX morrem porque há muita expressão do gene ligado ao X. Entretanto, em animais XY, mutações de perda de função no gene *Sxl* não têm efeito, o que é consistente com o fato de que a proteína SXL normalmente não é feita em machos.

As mutações de perda de função em *transformer* e *transformer2* têm o mesmo fenótipo: tanto animais XX quanto XY desenvolvem-se em machos. A transformação sexual em animais XX demonstra que ambos os genes *tra*<sup>+</sup> e *tra2*<sup>+</sup> são necessários para o desenvolvimento feminino. Entretanto, eles são perfeitamente dispensáveis para o desenvolvimento masculino. As mutações de perda de função no gene *dsx* fazem com que embriões XX e XY se desenvolvam em intersexos. O fenótipo





## ENFOQUE TÉCNICO: *Fruitless*

Em 1963, na reunião anual da American Society of Zoologists, Kulbir Gill relatou a descoberta de uma nova mutação de esterilidade em macho de *Drosophila*. Machos homozigotos para esta mutação não tentavam reproduzir-se com as fêmeas muito embora eles as cortejassem. Ainda mais marcante, os machos mutantes cortejavam outros machos, tanto mutantes quanto tipo selvagem. Este comportamento foi acentuadamente visto quando Gill colocou um grupo de machos mutantes juntos em uma cultura. Os machos formavam longas cadeias ou círculos, com cada macho na cadeia cortejando o macho na frente dele (Fig. 1). Entretanto, a mutação macho estéril aparentemente não tinha efeitos em fêmeas. As fêmeas homozigotas mutantes cruzavam prontamente com machos tipo selvagem e geravam uma prole.

Por muitos anos, a mutação macho estéril que Gill descobriu era pouco mais do que uma curiosidade de laboratório. Então, no final dos anos 1970 e começo de 1980, Jeffrey Hall e colaboradores fizeram experimentos para mapear a mutação e caracterizar seus efeitos fenotípicos mais completamente. A mutação também recebeu um nome, *fruitless (fru)*, em reconhecimento a seu efeito de esterilidade do macho.

Nos anos 1990, pesquisadores de vários outros laboratórios diferentes fizeram um esforço cooperativo para clonar o gene *fru*. Seu trabalho, publicado em 1996 com Lisa Ryner como autora principal,<sup>1</sup> estabeleceu que *fru* codifica um fator de transcrição tipo *zinc-finger* que regula os genes para comportamento sexual masculino – na verdade, a preferência sexual masculina – em *Drosophila*.

Os experimentos genéticos iniciais de Barbara Taylor sugeriram que *fru* controla uma ramificação na via de determinação do sexo em *Drosophila* (Fig. 22.9). As proteínas TRA e TRA2 regulam o processamento do RNA *dsx*. A recomposição alternativa de transcritos *dsx* em machos e fêmeas produz formas diferentes de um fator de transcrição (DSX-fêmea e DSX-macho) que regulam os genes especificamente envolvidos na diferenciação sexual. Entretanto, Taylor observou que a formação de um músculo abdominal particular em machos não era controlada pelo gene *dsx*. Este achado indicou que a via de determinação sexual tinha uma ramificação, regulada não por *dsx* mas por outro gene ainda não identificado, e que este gene desconhecido podia, como *dsx*, codificar fatores de transcrição específicos do sexo. Além disso, parecia provável que o gene desconhecido poderia ser regulado pós-transcricionalmente pelas proteínas de recomposição TRA e TRA2.

As proteínas TRA e TRA2 ligam-se a uma sequência de RNA com 13 nucleotídeos de tamanho. Essa sequência é repetida múltiplas vezes em transcritos do gene *dsx*. Assim, quando Ryner e seus colegas começaram seus esforços para clonar *fru*, eles procuraram outros genes que possuíam os sítios de ligação de RNA TRA/TRA2 com 13 nucleotídeos. Sua pesquisa os levou a um *locus* no



**Fig. 1** ■ Machos de *Drosophila* homozigotos para a mutação *fruitless (fru)* fazendo corte uns aos outros. Cada macho na cadeia corteja o macho a sua frente.

cromossomo 3 onde o gene *fru* foi mapeado. Clonagem e análise do DNA nesse *locus* revelou a estrutura molecular do gene *fru*. Ele contém três cópias da sequência de ligação ao RNA TRA/TRA2 com 13 nucleotídeos, e, como esperado, sua expressão é regulada ao nível da recomposição pelas proteínas TRA e TRA2.

O padrão da expressão de *fru* é complexo. Existem vários promotores, e os transcritos do gene são recompostos alternativamente em machos e fêmeas. Em machos, pelo menos um dos produtos de *fru* é um fator de transcrição. Não está claro se uma versão diferente deste fator de transcrição é feita em fêmeas. Experimentos de hibridização *in situ* com RNA indicam que a expressão de *fru* é limitada a um número relativamente pequeno de neurônios no sistema nervoso central da mosca, alguns dos quais foram implicados no controle do comportamento sexual masculino.

A descoberta do gene *fru* de *Drosophila* cria a possibilidade de que a orientação sexual esteja sob controle genético em outras espécies, incluindo a nossa. De fato, Dean Hamer e colaboradores sugeriram que um gene no braço longo do cromossomo X contribui para a homossexualidade em humanos.<sup>2</sup> Entretanto, a natureza deste gene, caso ele exista<sup>3</sup>, é desconhecida.

<sup>1</sup>Ryner, L. C., S. F. Goodwin, D. H. Castrillon, A. Anand, A. Vilella, B. S. Baker, J. C. Hall, B. J. Taylor e S. A. Wasserman. 1996. Control of male sexual behavior and sexual orientation in *Drosophila* by the *fruitless* gene. *Cell* 87: 1079-1089.

<sup>2</sup>Hamer, D. H., S. Hu, V. L. Magnuson, N. Hu e A. M. L. Pattatucci. 1993. A linkage between DNA markers on the X chromosome and male sexual orientation. *Science* 261:321-327.

<sup>3</sup>Rice, G., C. Anderson, N. Risch e G. Ebers. 1999. Male homosexuality: absence of linkage to microsatellite markers at Xq28. *Science* 284: 665-667.



intersexual aparece porque *ambas* as vias de desenvolvimento masculina e feminina são ativadas em mutantes *dsx*.

### Determinação do Sexo em *Caenorhabditis*

A via de determinação do sexo somático em *Caenorhabditis* envolve pelo menos 10 genes diferentes (Fig. 22.10). Como em drosófilas, mutações nestes genes alteram o desenvolvimento sexual (Quadro 22.1). Por exemplo, mutações de perda de função nos dois genes *transformer tra-1* e *tra-2* (não confundir com os genes *tra* e *tra2* de *Drosophila*) fazem com que animais XX se desenvolvam em machos e mutações de perda de função no gene *hermafrodite her-1* fazem com que animais XO se desenvolvam em hermafroditas. Estas alterações fenotípicas mostram que os produtos dos genes *tra-1* e *tra-2* são necessários para desenvolvimento hermafrodita normal e que o produto do gene *her-1* é necessário para desenvolvimento masculino normal. Mutações de perda de função em três outros genes, *fem-1*, *fem-2* e *fem-3* (para feminização), fazem com que animais XO se desenvolvam em fêmeas sem espermatozoides. Este fenótipo feminizado mostra que os produtos do gene *fem* são necessários para desenvolvimento masculino normal.

Pesquisadores foram capazes de arrumar estes seis genes em uma via de desenvolvimento. O primeiro gene, *her-1*, codifica uma proteína secretada que provavelmente é uma molécula sinalizadora. O gene seguinte, *tra-2*, codifica uma proteína ligada à membrana, que pode funcionar como um receptor para a proteína sinalizadora *her-1*. Os produtos dos

genes *fem* são proteínas citoplasmáticas que podem transduzir o sinal *her-1*, e o último gene na via, *tra-1*, aparentemente codifica um fator de transcrição tipo *zinc-finger*, que pode regular os genes envolvidos em diferenciação sexual.

Em *Caenorhabditis*, a via de determinação do sexo parece envolver uma série de reguladores negativos da expressão gênica. Em animais XO, o produto gênico secretado *her-1* aparentemente interage com o produto gênico *tra-2*, tornando-o inativo. Essa interação permite que os três produtos do gene *fem* sejam ativados, e eles inativam coletivamente o produto do gene *tra-1*, que é um regulador positivo de diferenciação feminina. Visto que o animal não pode desenvolver-se como hermafrodita sem a proteína ativa de *tra-1*, ele se desenvolve como macho. Em animais XX, a proteína *her-1* não é produzida. Assim, seu receptor hipotético, a proteína de *tra-2*, permanece ativo. A proteína ativa de *tra-2* faz com que o produto do gene *fem* seja inativado, o que por sua vez permite que a proteína de *tra-1* estimule a diferenciação feminina. Portanto, o animal desenvolve-se como hermafrodita.

O desenvolvimento sexual em *Caenorhabditis* depende fundamentalmente da proporção X:A, como ocorre em *Drosophila*. De algum modo, esta proporção é convertida em um sinal molecular que controla a diferenciação sexual. Este mesmo sinal também controla o fenômeno de compensação de dose, que em *Caenorhabditis* envolve hipotativação dos dois cromossomos X no hermafrodita (Cap. 21). O sinal da proporção X:A é direcionado para as vias de determinação do sexo e compensação de dose por meio de uma curta via envolvendo pelo menos quatro genes. Um deles, *xol-1*, é ne-

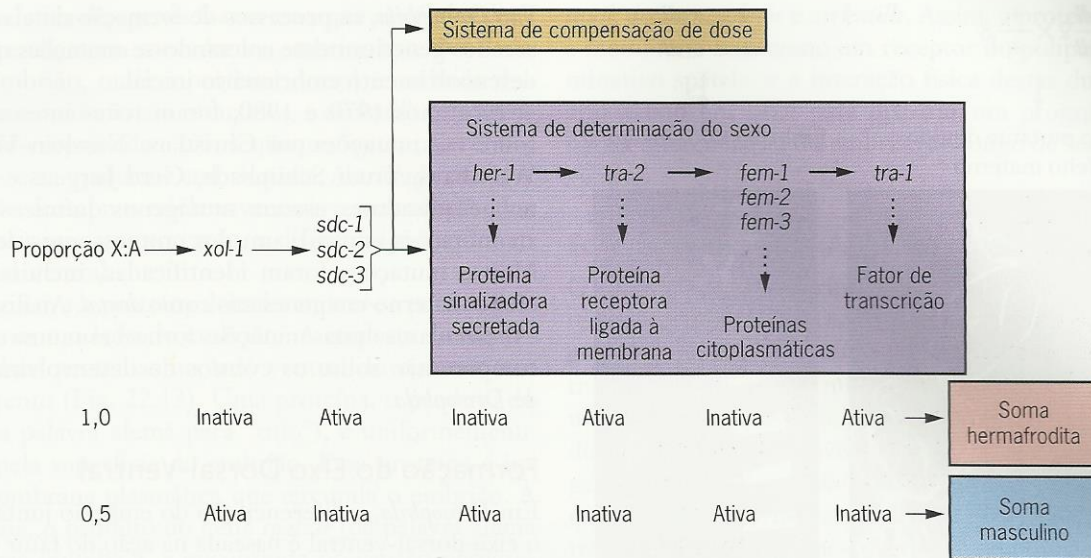
QUADRO 22.1

Fenótipos de Mutações de Perda de Função em Genes de Determinação de Sexo em *Drosophila melanogaster* e *Caenorhabditis elegans*<sup>a</sup>

| Gene                           | Fenótipo mutante XX                | Fenótipo mutante XY (ou XO) |
|--------------------------------|------------------------------------|-----------------------------|
| <i>Drosophila melanogaster</i> |                                    |                             |
| Gene numerador                 | letal                              | sem efeito                  |
| Gene denominador               | sem efeito                         | viabilidade reduzida        |
| <i>Sxl</i>                     | letal                              | sem efeito                  |
| <i>tra</i>                     | macho                              | sem efeito                  |
| <i>tra2</i>                    | macho                              | macho estéril               |
| <i>dsx</i>                     | intersexo estéril                  | intersexo estéril           |
| <i>Caenorhabditis elegans</i>  |                                    |                             |
| <i>xol-1</i>                   | sem efeito                         | letal                       |
| <i>sdc-1</i>                   | masculinizado                      | sem efeito                  |
| <i>sdc-2</i>                   | masculinizado                      | sem efeito                  |
| <i>sdc-3</i>                   | sem efeito na determinação do sexo | sem efeito                  |
| <i>her-1</i>                   | sem efeito                         | hermafrodita fértil         |
| <i>tra-2</i>                   | macho                              | sem efeito                  |
| <i>fem-1</i>                   | fêmea                              | fêmea                       |
| <i>fem-2</i>                   | fêmea                              | fêmea                       |
| <i>fem-3</i>                   | fêmea                              | fêmea                       |
| <i>tra-1</i>                   | macho                              | efeitos menores na gônada   |

<sup>a</sup>Fonte: Parkhurst, S. M., e P. M. Meneely. 1994. *Science* 264: 924-932.





**Fig. 22.10** ■ Via desenvolvimental para a determinação do sexo em *Caenorhabditis*. Diferenciação do animal como macho ou como hermafrodita depende de uma cascata de expressão gênica, que por sua vez depende da proporção X:A. A compensação de dose também depende da proporção X:A.

cessário em machos mas não em hermafroditas. Mutações de perda de função em *xol-1* fazem com que animais XO morram, donde o nome do gene, *XO-lethal*. Três outros genes, *sdc-1*, *sdc-2* e *sdc-3* (de determinação do sexo e compensação de dose), são regulados negativamente por *xol-1*. Mutações de perda de função nesses genes ou matam animais XX ou os transformam em machos. Assim, os genes *sdc* são necessários em hermafroditas mas não em machos.

### PONTOS IMPORTANTES

- Em *Drosophila*, a via que controla a diferenciação sexual envolve alguns genes que avaliam a proporção X:A, alguns que convertem esta proporção em um sinal de desenvolvimento e outros que respondem ao sinal produzindo estruturas masculinas ou femininas.
- O gene *Sex-lethal* (*Sxl*) tem um papel importante no desenvolvimento sexual de *Drosophila* regulando a recomposição de seu próprio transcrito e a de outro gene (*tra*).
- Em *C. elegans*, a via de diferenciação sexual envolve genes que codificam proteínas sinalizadoras, seus receptores, transdutores de sinal e fatores de transcrição.

## ATIVIDADE GÊNICA MATERNA NO DESENVOLVIMENTO

Os materiais transportados para o ovócito durante a ovocitogênese desempenham um papel importante no desenvolvimento embrionário.

Eventos importantes ocorrem no desenvolvimento animal mesmo antes do ovócito ser fertilizado. Nessa época,

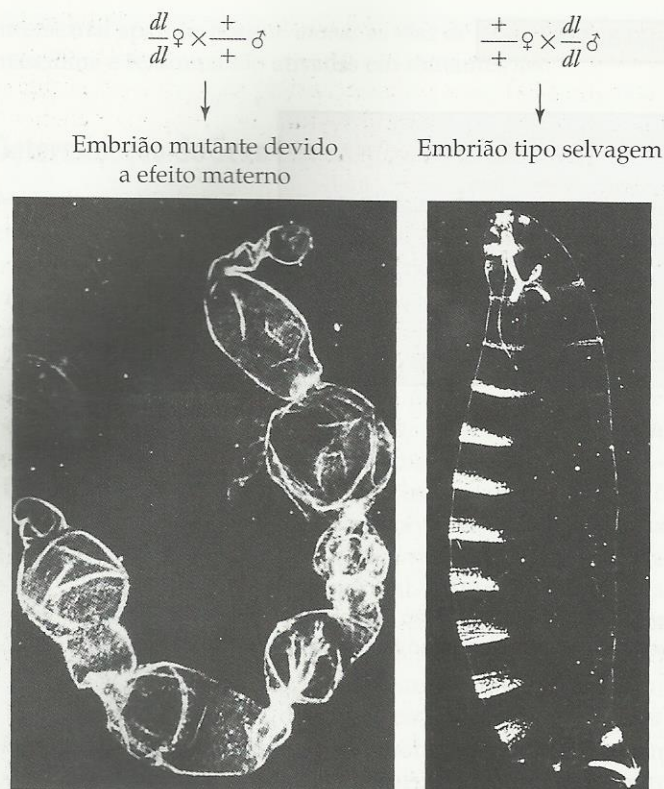
materiais nutritivos e determinativos são transportados para o ovócito vindos de células vizinhas, aumentando os estoques alimentícios e organizando o ovo para seu subsequente desenvolvimento, o equivalente molecular ao carinho materno. Esses materiais são gerados pela expressão de genes no sistema reprodutivo feminino, alguns sendo expressos em tecidos reprodutivos somáticos e outros apenas em tecidos da linhagem germinativa. Coletivamente, esses genes ajudam a formar ovos que podem desenvolver-se em embriões após a fertilização. Em algumas espécies, tais produtos gênicos maternos estabelecem o plano corpóreo básico do embrião, distinguindo a parte anterior da posterior e a parte dorsal da ventral. Estes materiais maternamente fornecidos portanto estabelecem um sistema molecular coordenado para guiar um desenvolvimento embrionário.

### Genes de Efeito Materno

Mutações em genes que contribuem para a formação de ovos saudáveis podem não ter efeito na viabilidade ou no aspecto das fêmeas que produzem esses ovos. Os efeitos só podem ser vistos na geração seguinte. Tais mutações são chamadas de **mutações de efeito materno** porque o fenótipo mutante na prole é causado por um genótipo mutante em sua mãe.

Os genes identificados por tais mutações são chamados de **genes de efeito materno**. O gene *dorsal* (*dl*) em *Drosophila* é um bom exemplo (Fig. 22.11). Reproduções entre moscas homozigotas para mutações recessivas neste gene produzem prole inviável. Este efeito letal é estritamente materno. Um cruzamento entre fêmeas mutantes homozigotas e machos tipo selvagem homozigotos produz prole inviável, mas o cruzamento recíproco (machos mutantes homozigotos × fêmeas tipo selvagem homozigotas) produz prole viável. O





**Fig. 22.11** ■ O efeito materno de uma mutação no gene *dorsal* (*dl*) de *Drosophila*. O fenótipo mutante é um embrião que não tem tecidos ventrais; isto é, é dorsalizado.

efeito letal da mutação *dorsal* portanto se manifesta apenas se as fêmeas forem homozigotas para ela. O genótipo do macho é irrelevante.

A caracterização molecular do gene *dorsal* revelou a base deste efeito materno. O gene *dorsal* codifica um fator de transcrição que é produzido durante a ovocitogênese e é estocado no ovo. No início do desenvolvimento, esse fator de transcrição tem um papel importante na diferenciação das partes dorsal e ventral do embrião. Quando ele está ausente, as partes ventrais diferenciam-se incorretamente como se estivessem na parte dorsal, criando um embrião com dois lados dorsais. Esta condição letal não pode ser evitada por um alelo *dorsal* tipo selvagem herdado do pai, pois *dorsal* não é transcrito no embrião. Como sua expressão é limitada à linhagem germinativa feminina, as mutações em *dorsal* são letais estritamente de efeito materno.

### Determinação dos Eixos Dorsal-Ventral e Anterior-Posterior em Embriões de *Drosophila*

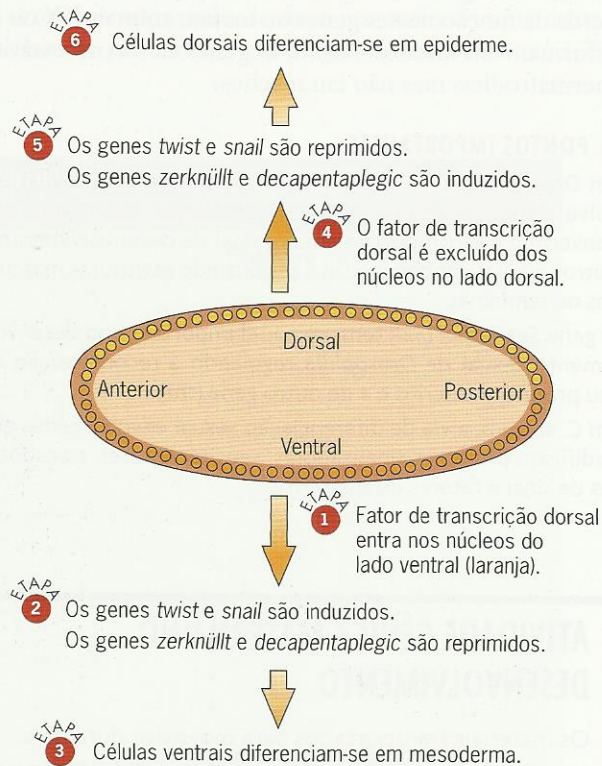
Animais com simetria bilateral têm dois eixos corpóreos primários, um que distingue o lado dorsal do ventral e o outro que distingue a cabeça da cauda (anterior do posterior). Ambos os eixos são estabelecidos muito cedo no desenvolvimento, em algumas espécies até mesmo antes da fertilização.

Em *Drosophila*, os processos de formação do eixo foram dissecados geneticamente coletando-se mutações que afetam o desenvolvimento embrionário inicial.

Nos anos 1970 e 1980, foram feitas intensas pesquisas sobre tais mutações por Christiane Nüsslein-Volhard, Eric Weischaus, Trudi Schüpbach, Gerd Jürgens e outros. Estes pesquisadores usaram mutágenos químicos para induzir mutações em cada um dos cromossomos de *Drosophila*. Muitas mutações foram identificadas, incluindo letais de efeito materno em genes tais como *dorsal*. Análises genéticas e moleculares destas mutações forneceram uma considerável compreensão sobre os eventos do desenvolvimento inicial de *Drosophila*.

### Formação do Eixo Dorsal-Ventral

Em *Drosophila*, a diferenciação do embrião juntamente com o eixo dorsal-ventral é baseada na ação do fator de transcrição codificado pelo gene *dorsal* (Fig. 22.12). Esta proteína é sintetizada maternamente e estocada no citoplasma do ovo. Na época da formação do blastodermis, a proteína dorsal entra no núcleo no lado ventral do embrião, induzindo a transcrição de dois genes chamados *twist* e *snail* (denominados por seus fenótipos mutantes). Nesses mesmos núcleos, ela reprime os genes *zerknüllt* (do alemão, enrugado) e *decapentaplegic* (das palavras gregas para “15” e “stroke”). A indução e



**Fig. 22.12** ■ Determinação do eixo dorsal-ventral em *Drosophila* pela proteína dorsal, um fator de transcrição que só atua nos núcleos do lado ventral do embrião. Os genes *twist*, *snail*, *zerknüllt* e *decapentaplegic* são regulados pela proteína dorsal.



a repressão seletiva dos genes fazem com que as células ventrais se diferenciem no mesoderma embrionário. No lado oposto do embrião, onde a proteína dorsal é excluída dos núcleos, *twist* e *snail* não são induzidos, e *zerknullt* e *decapentaplegic* não são reprimidos. Consequentemente, tais células diferenciam-se na epiderme embrionária. A entrada do fator de transcrição dorsal nos núcleos ventrais e sua exclusão dos núcleos dorsais iniciam, portanto, a diferenciação ao longo do eixo dorsal-ventral.

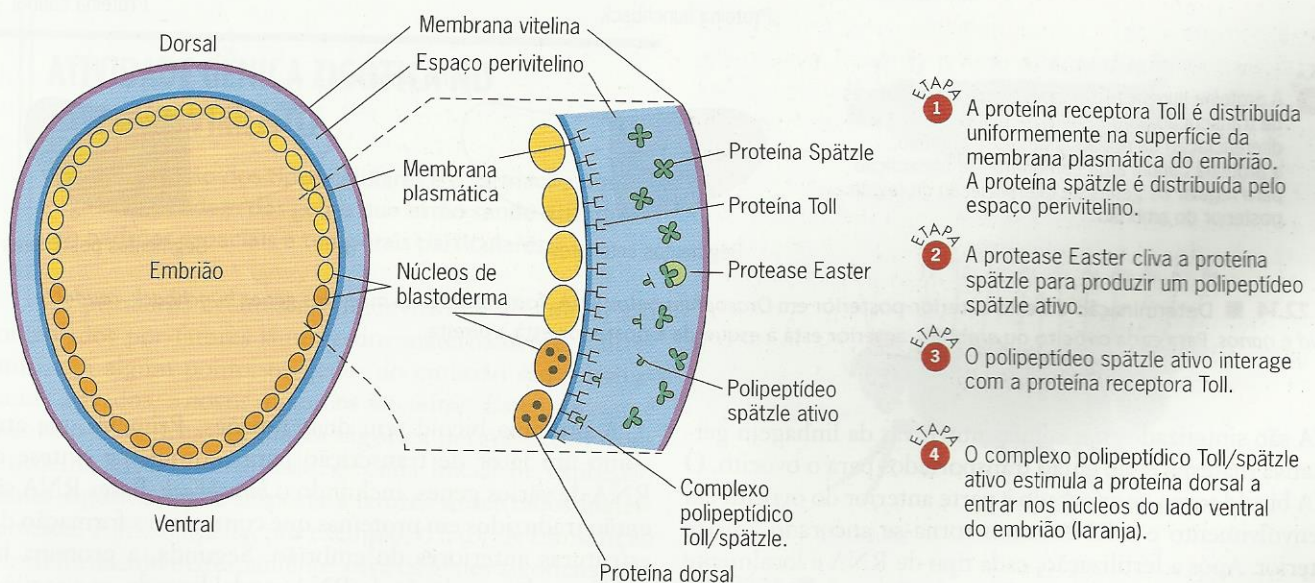
Mas o que ativa a proteína dorsal a mover-se para os núcleos de apenas um lado do embrião? A resposta é uma interação de duas proteínas no lado ventral do embrião em desenvolvimento (Fig. 22.13). Uma proteína, o produto do gene *Toll* (da palavra alemã para “tufo”), é uniformemente distribuída pela superfície do embrião. Essa proteína é inserida na membrana plasmática que circunda o embrião. A outra proteína, o produto do gene *spätzle* (da palavra alemã para “pequena bolinha”), é encontrada no espaço perivitelino, uma cavidade cheia de líquido entre a membrana plasmática e a membrana vitelina externa. Pela ação de uma protease codificada por um gene chamado *easter* (porque foi descoberto em Easter Sunday), a proteína *spätzle* é quebrada para produzir um polipeptídeo que interage com a proteína *Toll*. Entretanto, devido a um padrão estabelecido pelas células foliculares que circundam o ovócito antes dele ser fertilizado, a clivagem da proteína *spätzle* só ocorre no espaço perivitelínico do lado ventral do embrião. Quando a proteína *Toll* interage com o polipeptídeo *spätzle* gerado ventralmente, inicia-se uma cascata de eventos dentro do embrião que acaba por enviar a proteína dorsal para os núcleos embrionários. Lá, a proteína dorsal funciona como um fator de transcrição para regular a expressão dos genes *twist*,

*snail*, *decapentaplegic* e *zerknullt*. Assim, a proteína *Toll* ligada à membrana atua como um receptor do polipeptídeo determinativo *spätzle*, e a interação física destas duas moléculas atua como um sinal para disparar um programa genético para a diferenciação do embrião ao longo de seu eixo dorsal-ventral.

### Formação do Eixo Anterior-Posterior

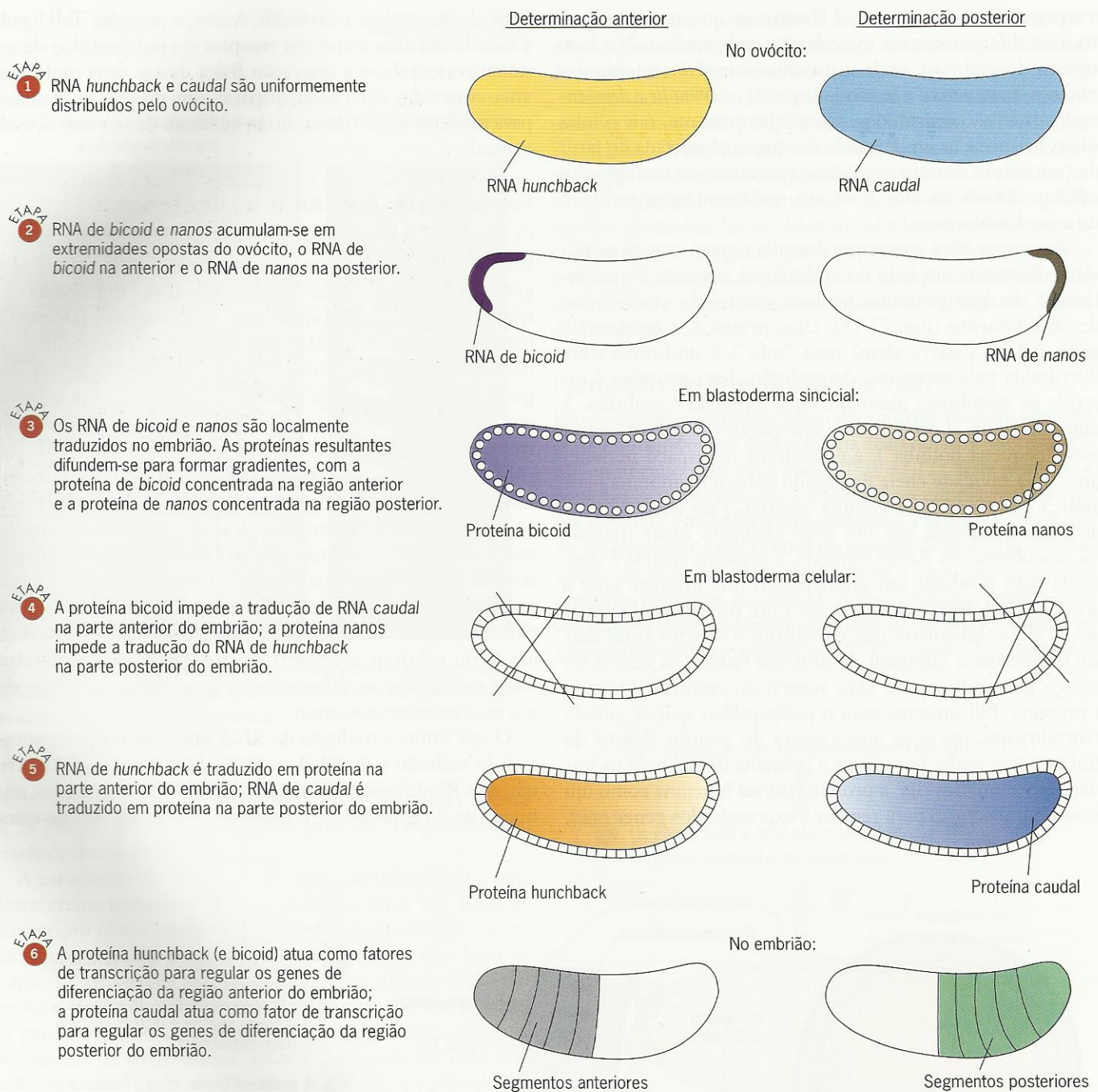
O eixo anterior-posterior em *Drosophila* é criado pela síntese regional de fatores de transcrição codificados pelos genes *hunchback* e *caudal* (Fig. 22.14). Estes dois genes são transcritos nas células nutritivas e na linhagem germinativa materna, isto é, as células que apóiam o crescimento e o desenvolvimento do ovócito. Os transcritos maternos dos genes *hunchback* e *caudal* são então levados para o ovócito, onde ficam uniformemente distribuídos no citoplasma. Entretanto, ambos os tipos de transcritos são traduzidos em partes diferentes do embrião. O RNA *hunchback* é traduzido apenas na parte anterior, e o RNA *caudal* é traduzido apenas na parte posterior. Esta tradução diferencial produz gradientes de concentração das proteínas codificadas por estes dois genes. A proteína *hunchback* é concentrada na parte anterior do embrião e a proteína *caudal* é concentrada na parte posterior. Estas duas proteínas então funcionam para ativar ou reprimir a transcrição dos genes cujos produtos estão envolvidos na diferenciação do embrião ao longo de seu eixo anterior-posterior.

O que limita a tradução do RNA *hunchback* à parte anterior do embrião e do RNA *caudal* à parte posterior? Ocorre que dois RNA fornecidos maternamente estão envolvidos, um transcrito do gene *bicoid* e o outro do gene *nanos*. Ambos estes



**Fig. 22.13** ■ Corte de um embrião de *Drosophila* mostrando a interação da proteína receptora Toll ligada à membrana com um polipeptídeo da proteína *spätzle* que induz a diferenciação ao longo do eixo dorsal-ventral. A formação do polipeptídeo *spätzle* de interação ocorre no espaço entre a membrana plasmática e a membrana vitelina do lado ventral do embrião.





**Fig. 22.14** ■ Determinação do eixo anterior-posterior em *Drosophila* pelos RNA, fornecidos pela mãe, dos genes *hunchback*, *caudal*, *bicoid* e *nanos*. Para cada ovócito ou embrião, anterior está à esquerda e posterior está à direita.

RNA são sintetizados nas células nutritivas da linhagem germinativa materna e são então transportados para o ovócito. O RNA *bicoid* torna-se ancorado à parte anterior do ovócito em desenvolvimento e o RNA *nanos* torna-se ancorado à parte posterior. Após a fertilização, cada tipo de RNA é localmente traduzido, e os produtos protéicos resultantes difundem-se pelo embrião para formar gradientes de concentração. A proteína *bicoid* é concentrada na extremidade anterior e a proteína *nanos* é concentrada na extremidade posterior.

A proteína *bicoid* tem duas funções. Primeira, ela atua como um fator de transcrição para estimular a síntese de RNA de vários genes, incluindo o *hunchback*. Esses RNA são então traduzidos em proteínas que controlam a formação das estruturas anteriores do embrião. Segunda, a proteína *bicoid* impede a tradução do RNA *caudal* ligando-se a sequências não traduzidas na região 3' deste RNA. Assim, sempre que a proteína *bicoid* é abundante (isto é, na parte anterior do embrião), o RNA *caudal* não é traduzido em proteína.



Contrariamente, sempre que a proteína bicoid é rara (isto é, na parte posterior do embrião), o RNA caudal é traduzido em proteínas. A regulação traducional do RNA caudal pela proteína bicoid é portanto responsável pelo gradiente de proteína caudal que se forma no embrião. Como a proteína caudal é um ativador específico de genes que controlam a diferenciação posterior, a parte do embrião que tem a mais alta concentração de proteína caudal desenvolve estruturas posteriores.

Ao contrário da proteína bicoid, a proteína nanos não funciona como um fator de transcrição. Entretanto, como a proteína bicoid, ela funciona como um regulador traducional. A proteína nanos é concentrada na parte posterior do embrião, e lá liga-se à região 3' não traduzida do RNA hunchback e causa a degradação deste RNA. Conseqüentemente, a proteína hunchback não é sintetizada na parte posterior do embrião. Em vez disso, sua síntese é restrita à parte anterior do embrião onde atua como um fator de transcrição para regular a expressão de genes envolvidos na diferenciação anterior-posterior. Sempre que a proteína hunchback é sintetizada, o embrião desenvolve estruturas anteriores.

#### PONTOS IMPORTANTES

- As proteínas e RNA codificados por genes de efeito materno tais como *dorsal*, *hunchback*, *bicoid* e *nanos* são transportados para óvulos de *Drosophila* durante a ovocitogênese.
- Produtos gênicos de efeito materno estão envolvidos na determinação dos eixos dorsal-ventral e anterior-posterior em embriões de *Drosophila*.
- As mutações recessivas em genes de efeito materno são expressas apenas em embriões produzidos por fêmeas homozigotas para estas mutações.

## ATIVIDADE GÊNICA ZIGÓTICA NO DESENVOLVIMENTO

A diferenciação dos tipos celulares e a formação dos órgãos dependem dos genes que estão sendo ativados em padrões espaciais e temporais particulares.

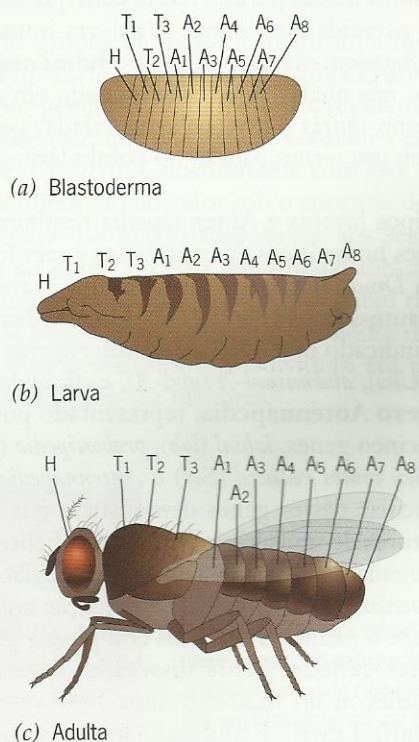
Os primeiros eventos no desenvolvimento animal são controlados por fatores sintetizados maternamente. Entretanto, em algum ponto, os genes do embrião são seletivamente ativados, e novos materiais são feitos. Este processo é chamado de **expressão gênica zigótica** porque ocorre após o ovócito ter sido fertilizado. A onda inicial de expressão gênica zigótica é uma resposta a fatores sintetizados maternamente. Em *Drosophila*, por exemplo, o fator de transcrição dorsal fornecido maternamente ativa os genes zigóticos *twist* e *snail*. À medida que o desenvolvimento continua, a ativação de outros genes zigóticos resulta em cascatas complexas de expressão gênica. Vamos agora examinar como estes genes zigóticos levam adiante o processo do desenvolvimento.

## Segmentação do Corpo

Em muitos invertebrados, o corpo consiste em um arranjo de unidades contíguas chamadas de **segmentos**. Uma *dro-sófila* adulta, por exemplo, tem uma cabeça, três segmentos torácicos distintos e oito segmentos abdominais. No tórax e no abdome, cada segmento pode ser identificado por coloração, padrão de cerdas e tipos de apêndices ligados a ele. Tais segmentos também podem ser identificados no embrião e na larva (Fig. 22.15). Em vertebrados, um padrão segmentar não é tão evidente no adulto, mas pode ser reconhecido no embrião pelo modo em que fibras nervosas crescem a partir do sistema nervoso central, a partir da formação dos arcos branquiais na cabeça e pela organização de massas musculares ao longo do eixo anterior-posterior. Mais adiante no desenvolvimento, estas características são modificadas, e o padrão segmentar original torna-se obscurecido. Entretanto, tanto em vertebrados quanto em muitos invertebrados, a segmentação é um aspecto fundamental do plano corpóreo geral.

## Genes Homeóticos

O interesse no controle genético da segmentação começou com a descoberta de mutações que transformam um segmento em outro. A primeira destas mutações foi encontrada



**Fig. 22.15** ■ Segmentação em *Drosophila* nos estágios de (a) blastoderma, (b) larval e (c) adulto do desenvolvimento. Embora os segmentos não estejam visíveis no blastoderma, suas células já estão programadas para formar os segmentos conforme mostrado; H, segmento da cabeça; T, segmento torácico; A, segmento abdominal.



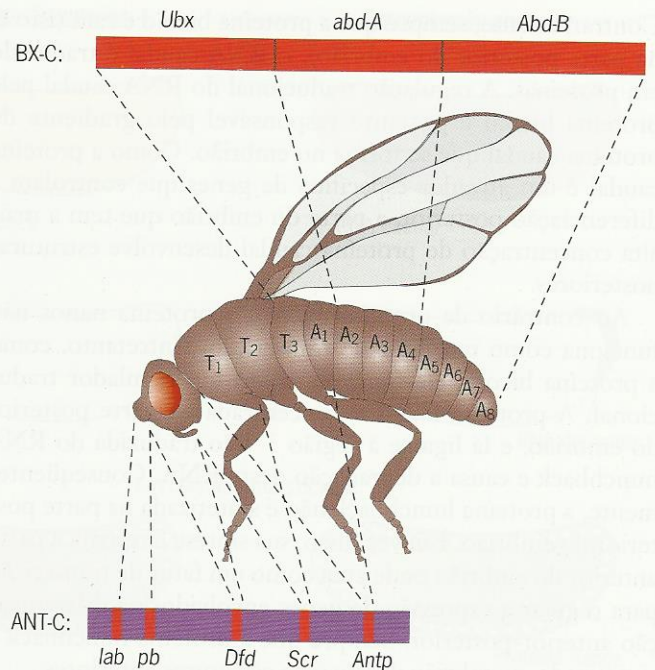


**Fig. 22.16** ■ O fenótipo de uma mutação *bithorax* em *Drosophila*.

em *Drosophila* em 1915 por Calvin Bridges. Ele a chamou de *bithorax* (*bx*) porque afeta dois segmentos torácicos. Neste mutante, o terceiro segmento torácico foi transformado, embora fracamente, no segundo, criando uma mosca com um pequeno par de asas rudimentares em lugar dos halteres (Fig. 22.16). Depois, outras mutações transformadoras de segmentos foram encontradas em *Drosophila*; por exemplo, *Antennapedia* (*Antp*), um mutante que transforma parcialmente as antenas na cabeça em pernas, que caracteristicamente crescem do tórax. Estas mutações foram chamadas de **mutações homeóticas** porque fazem com que uma parte do corpo fique parecida com outra. A palavra homeótica vem de William Bateson, que criou o termo **homeose** para referir-se a casos nos quais “algo havia mudado em outra coisa similar”. Como tantas outras palavras criadas por Bateson, esta se tornou um termo padrão no vocabulário da genética moderna.

Os fenótipos bitorax e Antennapedia resultam de mutações em genes homeóticos. Vários desses genes foram identificados em *Drosophila*, onde formam dois grandes grupos em um dos autossomos (Fig. 22.17). O **complexo bitorax**, geralmente indicado por **BX-C**, consiste em três genes, *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*), e *Abdominal-B* (*Abd-B*); o **complexo Antennapedia**, representado por **ANT-C**, consiste em cinco genes, *labial* (*lab*), *proboscipedia* (*pb*), *Deformed* (*Dfd*), *Sex combs reduced* (*Scr*) e *Antennapedia* (*Antp*). A análise molecular destes genes demonstrou que todos eles codificam fatores de transcrição hélice-giro-hélice com uma região conservada de 60 aminoácidos. Esta região, chamada de **homeodomínio**, está envolvida em ligação ao DNA.

O BX-C foi o primeiro dos dois complexos gênicos homeóticos a ser geneticamente dissecado. A análise deste complexo começou no final dos anos 1940 com o trabalho de Edward Lewis. Estudando mutações no BX-C, Lewis mostrou que a função tipo selvagem de cada parte do complexo é restrita a uma região específica no animal em desenvolvimento. Análises moleculares posteriores reforçaram e refinaram esta conclusão. O estudo de ANT-C começou nos anos 1970, principalmente pelo trabalho de



**Fig. 22.17** ■ Os genes homeóticos no complexo *bithorax* (BX-C) e no complexo *Antennapedia* (ANT-C) de *Drosophila*. São indicadas as regiões corpóreas nas quais cada gene se expressa.

Thomas Kaufman, Matthew Scott e seus colaboradores. Por uma combinação de análises genéticas e moleculares, tais pesquisadores mostraram que os genes do ANT-C também são expressos de um modo regionalmente específico. Entretanto, os genes ANT-C são expressos mais anteriormente que os genes BX-C. Curiosamente, o padrão de expressão dos genes ANT-C e BX-C juntamente com o eixo anterior-posterior corresponde exatamente à ordem dos genes ao longo do cromossomo (Fig. 22.17). Ainda não está claro por que é assim. A via desenvolvimental que cada célula toma parece depender simplesmente do conjunto de genes homeóticos que são expressos nela. Como os genes homeóticos têm tal papel central em selecionar as identidades segmentais de células individuais, eles geralmente são chamados de **genes seletores**.

As proteínas codificadas por genes homeóticos são fatores de transcrição homeodomínio. Essas proteínas ligam-se a seqüências reguladoras no DNA, incluindo algumas dentro dos próprios complexos bithorax e Antennapedia. Por exemplo, as proteínas UBX e ANTP ligam-se a uma seqüência dentro do promotor do gene *Ubx*, uma sugestão de que os genes homeóticos podem regular a si próprios e uns aos outros. Outros genes alvos dos fatores de transcrição homeodomínio foram identificados, incluindo alguns que codificam outros tipos de fatores de transcrição. Os genes homeóticos portanto parecem controlar uma cascata regulatória de genes alvos, que por sua vez atuam para determinar identidades segmentais de células individuais. Entretanto, os genes homeóticos não estão no topo desta cascata regu-



latória. Suas atividades são controladas por outro grupo de genes expressos mais cedo no desenvolvimento, os genes de segmentação.

### Genes de Segmentação

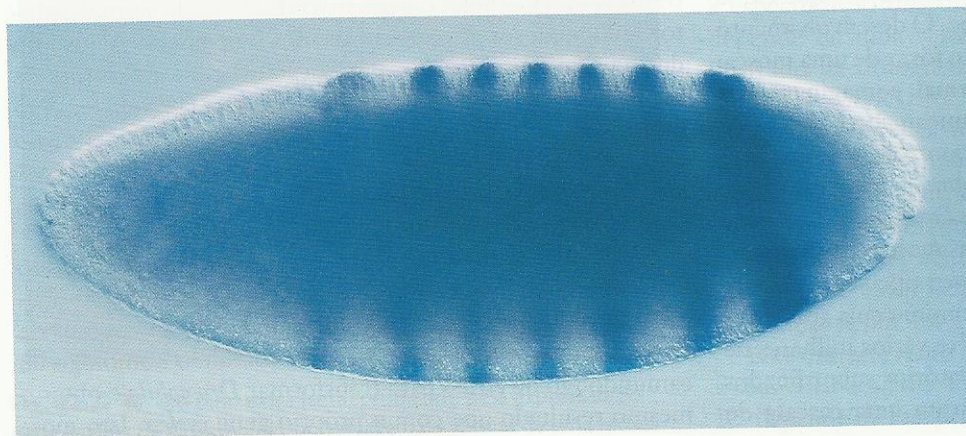
A maioria dos genes homeóticos são identificados por mutações que alteram o fenótipo da mosca adulta. Entretanto, essas mesmas mutações também têm efeitos fenotípicos nos estágios embrionários e larvais. Este achado sugeriu que outros genes envolvidos em segmentação poderiam ser descobertos triando-se mutações causadoras de defeitos embrionários e larvais. Nos anos 1970 e 1980, Christiane Nüsslein-Volhard e Eric Wieschaus fizeram tais triagens (veja Um Marco na Genética: Mutações que Perturbam a Segmentação em *Drosophila*, no final deste capítulo). Eles observaram um conjunto novo de genes necessários para segmentação ao longo do eixo anterior-posterior. Nüsslein-Volhard e Wieschaus classificaram estes **genes de segmentação** em três grupos com base em fenótipos mutantes embrionários.

1. **Genes Gap.** Estes genes definem regiões segmentares no embrião. Mutações nos genes *gap* causam ausência de todo um grupo de segmentos corpóreos contíguos; isto é, eles criam um espaço anatômico ao longo do eixo anterior-posterior. Quatro genes *gap* foram bem caracterizados: *Krüppel* (da palavra em alemão para “aleijado”), *giant*, *bunchback* e *knirps* (do alemão para “anão”). Cada um é expresso em regiões características no embrião inicial sob controle dos genes de efeito materno *bicoid* e *nanos*. Os genes *gap* codificam fatores de transcrição.
2. **Genes Pair-rule.** Estes genes definem um padrão de segmentos dentro do embrião. Os genes *pair-rule* são regulados pelos genes *gap* e são expressos em sete bandas, ou faixas, alternadas ao longo do eixo anterior-posterior, dividindo o embrião em 14 zonas distintas, ou **parassegmentos** (Fig. 22.18). Mutações em cada um dos vários genes *pair-rule* produzem embriões com apenas metade dos muitos parasseg-

mentos que o tipo selvagem tem. Em cada mutante, há parassegmentos faltando, embora estes não sejam os mesmos em diferentes mutantes *pair-rule*. Exemplos de genes *pair-rule* são *fushi tarazu* (do japonês para “algo faltando”) e *even-skipped*. Em mutantes *fushi tarazu*, cada um dos parassegmentos de número ímpar está faltando; em mutantes *even-skipped*, cada um dos parassegmentos de número par está faltando. A maioria dos genes *pair-rule* codificam fatores de transcrição.

3. **Genes Segment-Polarity.** Estes genes definem os compartimentos anterior e posterior de segmentos individuais ao longo do eixo anterior-posterior. Mutações nos genes de *segment-polarity* fazem com que parte de cada segmento seja substituída por uma cópia em imagem especular de um meio-segmento adjacente. Por exemplo, mutações no gene *segment-polarity gooseberry* fazem com que a metade posterior de cada segmento seja substituída por uma cópia em imagem especular do meio-segmento anterior adjacente. Muitos dos genes *segment-polarity* são expressos em 14 bandas estreitas ao longo do eixo anterior-posterior. Assim, eles refinam o padrão segmental estabelecido pelos genes *pair-rule*. Dois dos genes *segment-polarity* mais bem estudados são o *engrailed* e *wingless*. *Engrailed* codifica um fator de transcrição e *wingless* codifica uma molécula sinalizadora.

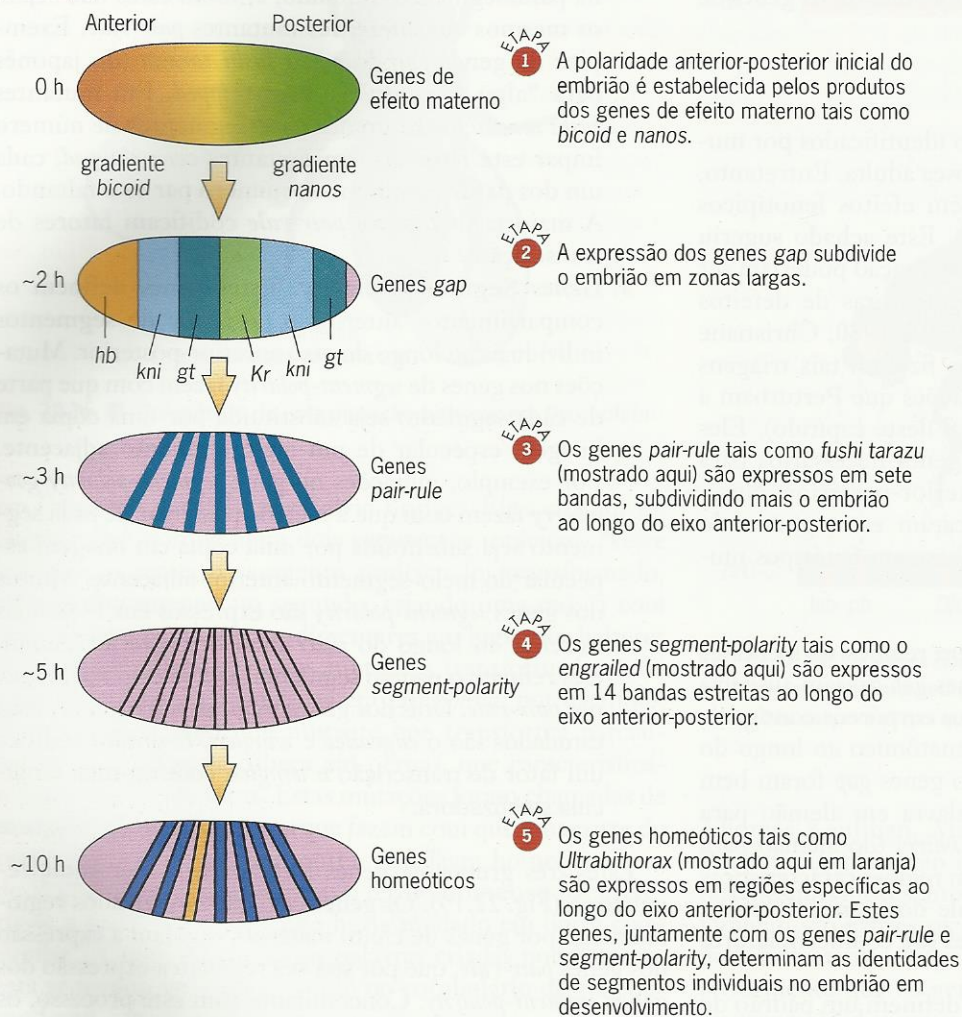
Estes três grupos de genes formam uma hierarquia regulatória (Fig. 22.19). Os genes *gap*, que são ativados regionalmente por genes de efeito materno, regulam a expressão dos genes *pair-rule*, que por sua vez regulam a expressão dos genes *segment-polarity*. Concomitante com este processo, os genes homeóticos são ativados sob o controle dos genes *gap* e *pair-rule* para dar identidades únicas aos segmentos que se formam ao longo do eixo anterior-posterior. Interações entre os produtos de todos estes genes então refinam e estabilizam os limites segmentais. Deste modo, o embrião de *Drosophila* é progressivamente subdividido em unidades desenvolvimentais cada vez menores.



**Fig. 22.18 ■** O padrão *seven-stripe* (sete listras) de expressão do RNA do gene *pair-rule fushi tarazu (ftz)* em um blastoderm de embrião de *Drosophila*. O RNA foi detectado por hibridização *in situ* com uma sonda específica de *ftz*. Anterior é à esquerda e dorsal em cima. Outros genes *pair-rule* mostram um padrão *seven-stripe* diferente.



Horas após  
fertilização



**Fig. 22.19** ■ Cascata de expressão gênica para produzir segmentação em embriões de *Drosophila*.

## Formação de Órgão

Quando muitos tipos diferentes de células são organizadas para um fim específico, elas formam um órgão. O coração, o estômago, o rim, o fígado e o olho são exemplos de órgãos. Uma das marcantes características de um órgão é que ele se forma em uma parte específica do corpo. O desenvolvimento de um coração na cabeça ou um olho no tórax de uma mosca, por exemplo, seria extremamente anormal, e imaginá-los o que deu errado. A formação anatomicamente correta de um órgão está obviamente sob rígido controle genético.

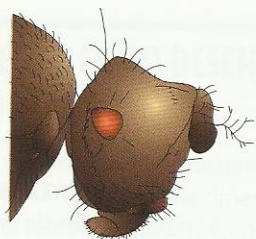
Geneticistas obtiveram *insights* sobre a natureza deste controle pelo estudo de outro gene em *Drosophila*. Este gene é chamado *eyeless* devido ao fenótipo das moscas que são mutantes para ele (Fig. 22.20). O gene *eyeless* tipo selvagem codifica um fator de transcrição homeodomínio cuja ação ativa uma via de desenvolvimento que envolve vários milhares de genes. Inicialmente, vários genes regulatórios subordinados são ativados. Seus produtos então ativam uma cascata de

eventos que criam tipos celulares específicos dentro do olho em desenvolvimento.

O papel do gene *eyeless* foi demonstrado por sua expressão em tecidos que normalmente não formam olhos (Fig. 22.21). Walter Gehring e colaboradores o fizeram criando moscas transgênicas nas quais o gene *eyeless* foi fundido a um promotor que podia ser ativado em tecidos específicos. A ativação desse promotor causou a transcrição do gene *eyeless* fora de seu domínio normal de expressão. Isto, por sua vez, fez com que os olhos se formassem em locais estranhos tais como asas, pernas e antenas. Esses olhos extras (ou ectópicos) eram anatomicamente bem desenvolvidos e funcionais. De fato, seus fotorreceptores respondiam a luz.

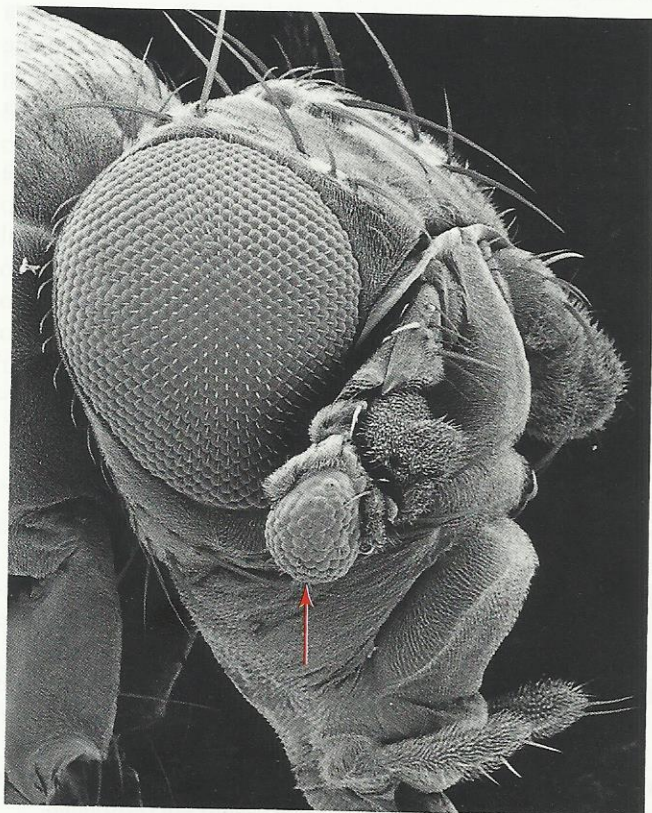
Um achado ainda mais marcante é que um homólogo em mamíferos do gene *eyeless*, chamado *Pax6*, também produz esses olhos extras quando inserido em cromossomos de *Drosophila*. Gehring e colaboradores usaram o homólogo de camundongo de *eyeless* para transformar *Drosophila* e obter o mesmo resultado que com o próprio gene *eyeless*. Isto mos-





**Fig. 22.20** ■ O fenótipo de um mutante *eyeless* em *Drosophila*.

trou que o gene de camundongo, que também codifica uma proteína homeodomínio, é funcionalmente equivalente ao gene de *Drosophila*. Isto é, ele regula a via para o desenvolvimento do olho. Entretanto, quando o gene de camundongo é colocado em *Drosophila*, ele produz olhos de *Drosophila*, e não de camundongo. Olhos de *Drosophila* desenvolvem-se porque os genes que respondem a este comando regulador do gene de camundongo inserido são genes normais de *Drosophila*, que devem, lógico, especificar a formação de um olho de *Drosophila*. Em camundongo, mutações no homólogo do gene *eyeless* reduzem o tamanho dos olhos. Por este motivo, o fenótipo mutante é chamado *Small eye*. Um homólogo de *eyeless* e *Small eye* também foi encontrado em seres humanos. Mutações neste gene causam uma síndrome de defeitos oculares chamada *aniridia*, na qual a íris está reduzida ou ausente.



**Fig. 22.21** ■ Um olho extra (seta) produzido pela expressão do gene *eyeless* de *Drosophila* tipo selvagem na antena de uma mosca.

A descoberta de genes homólogos que controlam o desenvolvimento do olho em organismos diferentes tem profundas implicações evolutivas. Ela sugere que a função destes genes é muito antiga, vindo do ancestral comum de moscas e mamíferos. Talvez os olhos neste organismo ancestral não fossem mais do que um grupo de células sensíveis a luz organizadas pelos efeitos reguladores de um gene *eyeless* primitivo. Com o tempo evolutivo, este gene continuou a regular o processo cada vez mais complicado de desenvolvimento do olho, de modo que hoje olhos tão diferentes como os de insetos e de mamíferos ainda são formados sob seu controle.

### Especificação de Tipos Celulares

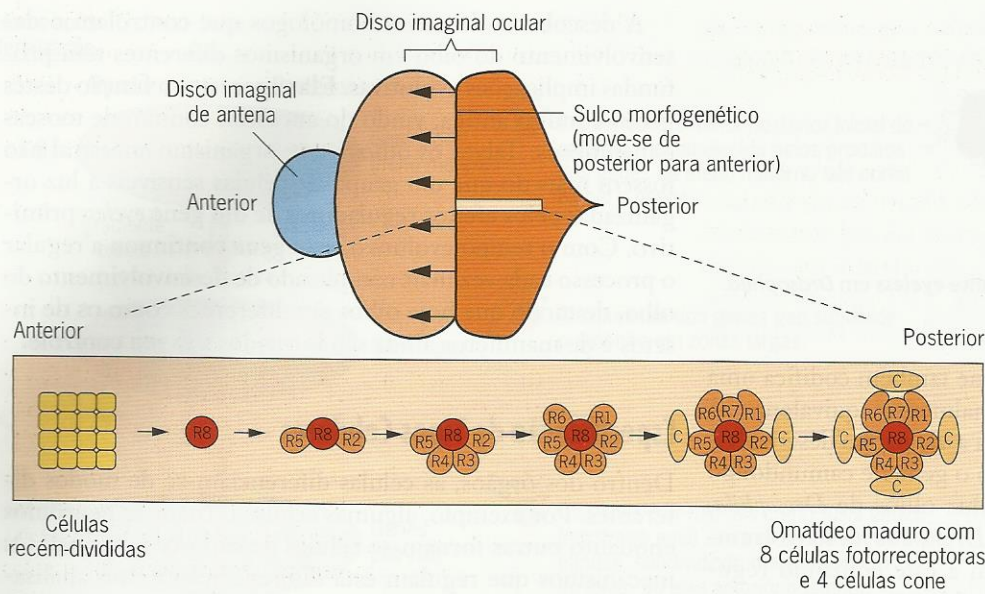
Dentro dos órgãos, as células diferenciam-se de modos diferentes. Por exemplo, algumas células tornam-se neurônios enquanto outras tornam-se células de suporte neuronal. Os mecanismos que regulam esta diferenciação foram analisados estudando-se situações muito simples que envolvem alguns tipos celulares distintos. Uma destas situações ocorre no desenvolvimento do olho de *Drosophila* (Fig. 22.22).

Cada um dos grandes olhos compostos de *Drosophila* origina-se como uma camada achatada de células em um dos discos imaginais. Inicialmente, todas as células nesta camada epitelial parecem iguais, porém mais tarde durante o estágio larval, forma-se um sulco perto da margem posterior do disco. À medida que esse sulco se move na direção anterior através do disco, ele ativa uma onda de divisões celulares em seu caminho. As células recém-divididas então se diferenciam em tipos celulares específicos para formar as 800 facetas individuais do olho adulto. Cada faceta consiste em 20 células. Oito são neurônios fotorreceptores destinados a absorver luz; quatro são células cone que secretam uma lente para enforçar a luz nos fotorreceptores; seis são células de bainha para dar delimitação e suporte; e as duas células restantes formam pêlos sensoriais na superfície do olho. Assim, um arranjo altamente padronizado de complexas facetas diferenciadas desenvolve-se a partir do que foi uma camada achatada de células idênticas. O que causa esta transformação?

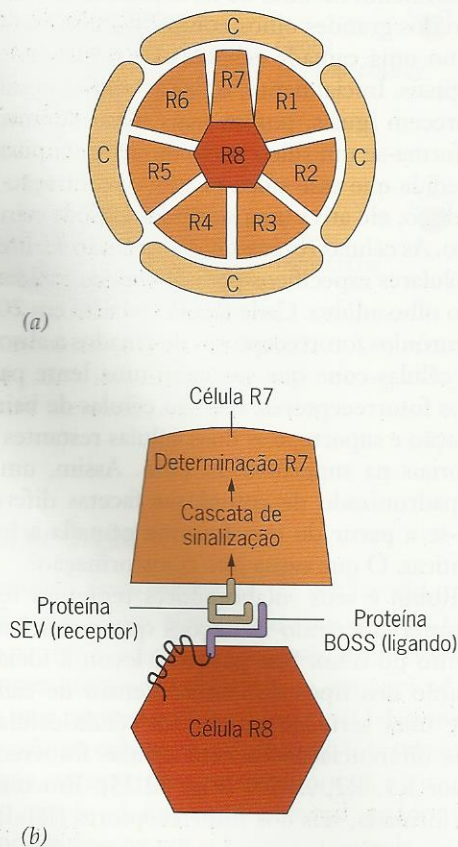
Gerald Rubin e seus colaboradores tentaram responder a esta pergunta coletando mutações que perturbam o desenvolvimento do olho. Sua pesquisa levou à idéia de que a especificação dos tipos de células dentro de cada faceta depende de uma série de interações célula-célula. Isto é ilustrado na diferenciação de oito células fotorreceptoras, indicadas por R1, R2, ... R8 (Fig. 22.23). Em uma faceta totalmente formada, seis dos fotorreceptores (R1-R6) estão dispostos em um círculo ao redor dos outros dois (R7, R8). Uma das células centrais, R8, é a primeira a se diferenciar na faceta em desenvolvimento. Seu aparecimento é seguido da diferenciação das células periféricas R2 e R5, então de R3 e R4 e R1 e R6. Finalmente, a segunda célula central, R7, diferencia-se em um fotorreceptor.

Este último evento foi estudado em grande detalhe. Rubin e seus colaboradores mostraram que a diferenciação da célula R7 depende da recepção de um sinal da célula R8 já diferen-





**Fig. 22.22** ■ Desenvolvimento do olho de *Drosophila*. À medida que o sulco morfogênico se move para o anterior do disco imaginal olho-antena, ocorre uma onda de divisões celulares. As células recém-divididas então começam a se diferenciar em tipos específicos. A inserção mostra a diferenciação das células de fotorreceptores (R1–R8) e cones (C) que formam cada omatídeo (faceta) do olho composto.



**Fig. 22.23** ■ Determinação do fotorreceptor R7 em um omatídeo (faceta) no olho composto de *Drosophila*. (a) Disposição de oito fotorreceptores (1–8) e quatro células cone (C) em um omatídeo. (b) Sinalização entre a célula R8 diferenciada e a suposta célula R7. A proteína *bride of sevenless* (BOSS) na célula R8 é o ligando para a proteína receptora *sevenless* (SEV) na superfície da célula R7. A ativação deste receptor inicia uma cascata de sinalização dentro da célula R7 que a induz a se diferenciar.

ciada. Para receber este sinal, a célula R7 deve produzir um receptor específico, uma proteína ligada à membrana codificada por um gene chamado *sevenless* (*sev*). Mutações neste gene abolem a função do receptor e impedem que a célula R7 se diferencie como um neurônio. Em vez disso, ela se diferencia como uma célula cone. O sinal para o receptor R7 é produzido por um gene chamado *bride of sevenless* (*boss*) e é especificamente expresso na superfície da célula R8. O contato entre a célula diferenciada R8 e a célula indiferenciada R7 permite que o sinal R8, ou **ligando** como é tecnicamente chamado, interaja com o receptor R7 e o ative. Essa ativação induz uma cascata de mudanças dentro da célula R7 que no final a leva a diferenciar-se como um neurônio receptor de luz. Tal diferenciação é supostamente mediada por um ou mais fatores de transcrição que agem em genes dentro do núcleo de R7. Assim, o sinal da célula R8 é “transduzido” no núcleo de R7, onde altera o padrão de expressão gênica. A análise do desenvolvimento do olho em *Drosophila* portanto mostra que a **indução**, o processo de determinação do destino de uma célula indiferenciada por um sinal de uma célula diferenciada, pode ter um papel importante na especificação dos tipos celulares.

#### ■ PONTOS IMPORTANTES

- Os genes zigóticos são ativados após fertilização em resposta a produtos gênicos maternos.
- Em *Drosophila*, os produtos dos genes de segmentação regulam a subdivisão do embrião em uma série de segmentos ao longo do eixo anterior-posterior.
- A identidade de cada segmento do corpo é determinada pelos produtos dos genes nos complexos gênicos homeóticos *hox* e *Antennapedia*.
- A formação de um órgão pode depender do produto de um gene master regulatório, tal como o gene *eyeless* em *Drosophila*.
- Em *Drosophila*, tipos de genes específicos diferenciam-se após terem sido estabelecidas as identidades segmentais.



## ANÁLISE GENÉTICA DO DESENVOLVIMENTO EM VERTEBRADOS

Geneticistas podem estudar o desenvolvimento em vertebrados aplicando os conhecimentos obtidos pelo estudo de invertebrados modelos, analisando mutações e fenocópias de genes mutantes em vertebrados modelos tais como camundongos e peixe-zebra e examinando a diferenciação de células tronco.

Grande parte do conhecimento sobre o controle genético do desenvolvimento veio do estudo de *Drosophila* e *Caenorhabditis*, dois invertebrados modelos. Os geneticistas gostariam de aplicar e estender este conhecimento a outros grupos de animais – em particular, vertebrados. A meta final seria aprender sobre o controle genético do desenvolvimento em nossa própria espécie. Uma estratégia para atingir esta meta é usar a informação obtida do estudo de genes de invertebrados para identificar genes significativos do ponto de vista do desenvolvimento em vertebrados. Outra é estudar espécies de vertebrados modelos com técnicas similares às que estão sendo usadas em *Drosophila* e *Caenorhabditis*.

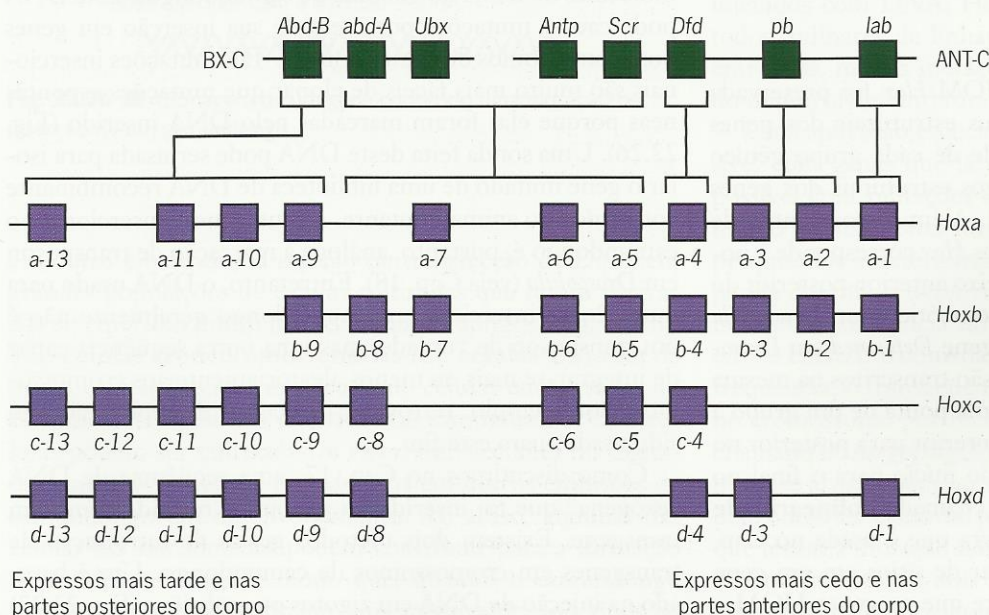
### Homólogos em Vertebrados de Genes de Invertebrados

Uma vez que um gene tenha sido isolado e seqüenciado, pesquisadores podem triar bancos de dados de seqüências de DNA quanto a genes homólogos em outros organismos. Se as seqüências de genes tiverem sido razoavelmente bem conservadas durante o tempo evolutivo, este procedimento funciona mesmo para espécies distantemente relacionadas.

Assim, foi possível identificar genes de várias espécies de vertebrados que são homólogos a genes de *Drosophila* e *C. elegans*. A identificação de um gene de vertebrado possibilita muitos tipos de análises experimentais, incluindo análises da expressão gênica tanto no que concerne a RNA quanto a proteínas.

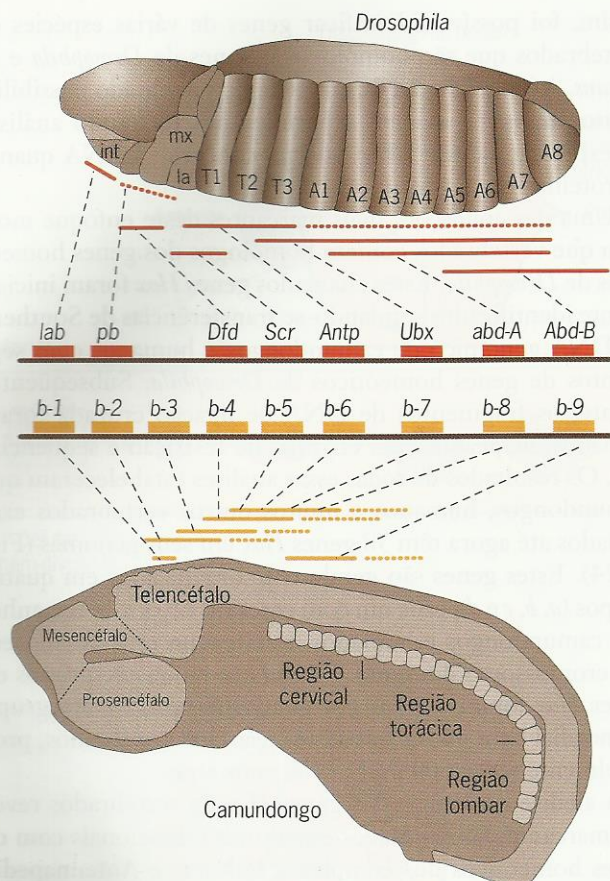
Uma das aplicações mais marcantes deste enfoque mostrou que vertebrados contêm homólogos dos genes homeóticos de *Drosophila*. Estes chamados genes *Hox* foram inicialmente identificados sondando-se transferências de Southern de DNA genômico de camundongos e humanos com segmentos de genes homeóticos de *Drosophila*. Subseqüentemente, os fragmentos de DNA de ligação cruzada foram clonados, mapeados com enzimas de restrição e seqüenciados. Os resultados de todas essas análises estabeleceram que camundongos, humanos e muitos outros vertebrados examinados até agora têm 38 genes *Hox* em seus genomas (Fig. 22.24). Estes genes são geralmente organizados em quatro grupos (*a*, *b*, *c* e *d*), cada um com cerca de 120 kb de tamanho. Em camundongos e humanos, cada grupo está situado em um cromossomo diferente. Parece que os quatro grupos de genes *Hox* foram criados por quadruplicação de um grupo primordial bem no início da evolução dos vertebrados, provavelmente 500 a 600 milhões de anos atrás.

A análise molecular dos genes *Hox* de vertebrados revelou marcantes similaridades estruturais e funcionais com os genes homeóticos dos complexos bithorax e Antennapedia em *Drosophila*. Estes últimos genes na verdade parecem ter sido membros de um grupo maior, chamado HOM-C, que foi dividido por um rearranjo cromossômico durante a evolução das moscas. Em outros tipos de insetos tais como o besouro da farinha, *Tribolium castaneum*, os complexos bithorax e Antennapedia são unidos em um único grupo. A comparação entre vertebrados e invertebrados mostra que



**Fig. 22.24** ■ Organização e expressão dos genes *Hox* de mamíferos homólogos aos genes de *Drosophila* em BX-C e ANT-C. As homologias são indicadas por colchetes. Todos os genes exceto *Dfd* são transcritos da esquerda para a direita. São indicados a época e o local anatômico de expressão.





**Fig. 22.25** ■ Padrões de expressão de genes homeóticos de *Drosophila* e camundongo. Em cima: embrião de *Drosophila* com 10 horas de idade mostrando a expressão aproximada dos genes homeóticos na epiderme. Na cabeça, int, mx e la designam segmentos intercalar, maxilar e labial, respectivamente. Os segmentos torácico e abdominal são indicados como T1–T3 e A1–A8. Embaixo: um embrião de camundongo com 12 dias mostrando a expressão aproximada dos genes *Hoxb* no sistema nervoso central. As linhas tracejadas indicam a extensão da expressão para o lado posterior.

a organização básica dos genes HOM/*Hox* foi preservada durante a evolução. Os homólogos estruturais dos genes ANT-C estão em uma extremidade de cada grupo gênico *Hox* de vertebrados, e os homólogos estruturais dos genes BX-C estão na outra extremidade. Além disso, dentro de cada grupo, a ordem física dos genes *Hox* corresponde à posição de sua expressão ao longo do eixo anterior-posterior do embrião, como ocorre nos genes homeóticos em *Drosophila* (Fig. 22.25). Com uma exceção (o gene *Deformed* em *Drosophila*), todos os genes HOM e *Hox* são transcritos na mesma direção, com a expressão indo de uma ponta de um grupo à outra ponta, tanto espacialmente (anterior para posterior no embrião) quanto temporalmente (do início para o final no desenvolvimento). Este fenômeno, chamado **colinearidade** (não confundir com a mesma palavra que é usada no Cap. 14 para descrever a sequência linear de sítios em um gene e seu produto polipeptídico), sugere que os genes HOM e

*Hox* fornecem um mecanismo molecular comum para estabelecer as identidades de regiões específicas em muitos tipos diferentes de embriões.

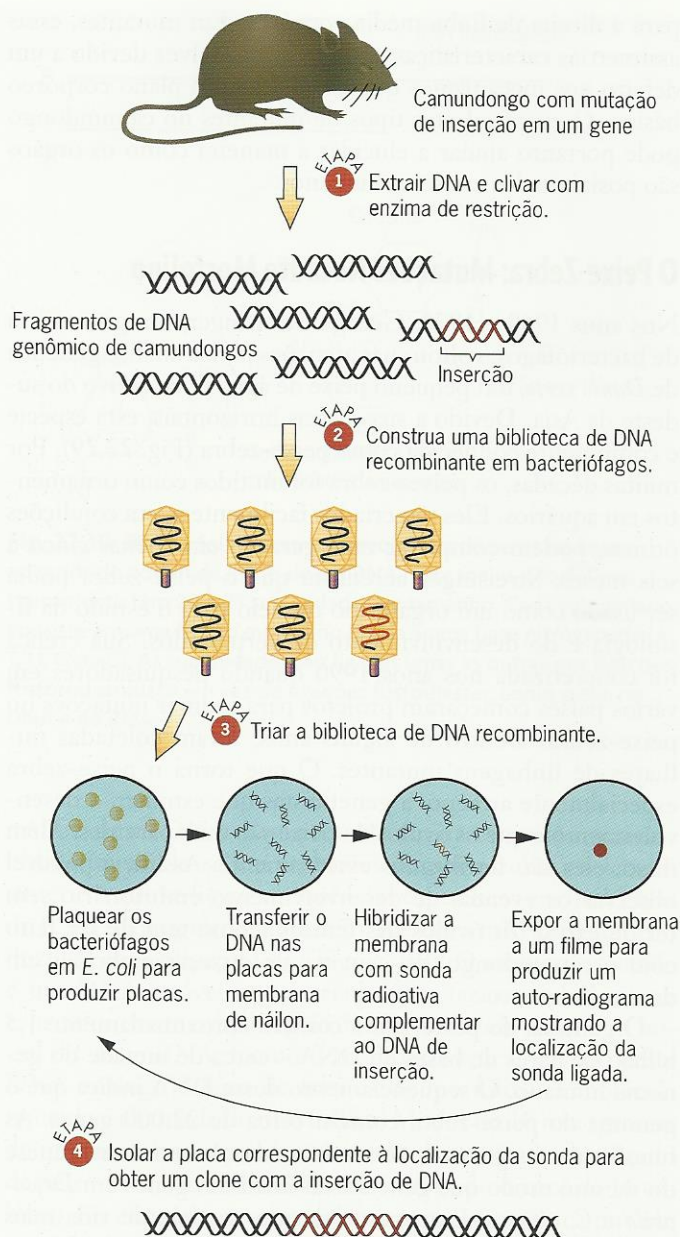
### O Camundongo: Mutagênese Insercional, Transgênicos e Mutações Nocaute

O controle genético do desenvolvimento não pode ser estudado em vertebrados com a mesma precisão com que é em invertebrados tais como *Drosophila*. Obviamente, existem restrições técnicas e logísticas. Vertebrados têm ciclos de vida comparativamente mais longos, sua criação é cara e é difícil obter e analisar linhagens mutantes, especialmente aquelas com um significado desenvolvimental. A despeito destas restrições, geneticistas foram capazes de fazer avanços na análise genética do desenvolvimento em um vertebrado complexo, o camundongo, e estão sendo feitos esforços para analisar o desenvolvimento em outro vertebrado, o peixe-zebra. Mais de 500 *loci* responsáveis por doenças genéticas foram identificados no camundongo, e alguns estão envolvidos em processos de desenvolvimento. A maioria desses *loci* foram descobertos por projetos continuados para coletar mutações espontâneas. Tal trabalho requer que números muito grandes de camundongos, centenas de milhares, ou mesmo milhões, sejam criados e examinados quanto a diferenças fenotípicas e que qualquer diferença encontrada seja testada quanto a transmissão genética. Este trabalho árduo e caro só pode ser feito em algumas instalações em todo o mundo. Uma vez que seja detectada uma mutação, ela pode ser mapeada e testada quanto a alelismo e interações epistáticas com outras mutações. Entretanto, a clonagem do gene mutante pode demonstrar-se tão difícil e cara quanto encontrar o camundongo mutante original. Felizmente, a descoberta de novas técnicas acelerou este processo.

Tais técnicas são baseadas na introdução de determinadas moléculas de DNA no genoma de camundongos. Este DNA pode causar mutações por meio de sua inserção em genes nos cromossomos de camundongos. Tais mutações insercionais são muito mais fáceis de clonar que mutações espontâneas porque elas foram marcadas pelo DNA inserido (Fig. 22.26). Uma sonda feita deste DNA pode ser usada para isolar o gene mutado de uma biblioteca de DNA recombinante construída do animal mutante. A mutagênese insercional no camundongo é, portanto, análoga à marcação de transposon em *Drosophila* (veja Cap. 18). Entretanto, o DNA usado para mutagênese insercional no camundongo geralmente não é um transposon de verdade, mas uma outra sequência capaz de integrar-se mais ou menos aleatoriamente aos cromossomos. Por exemplo, retrovírus defectivos de replicação têm sido usados para este fim.

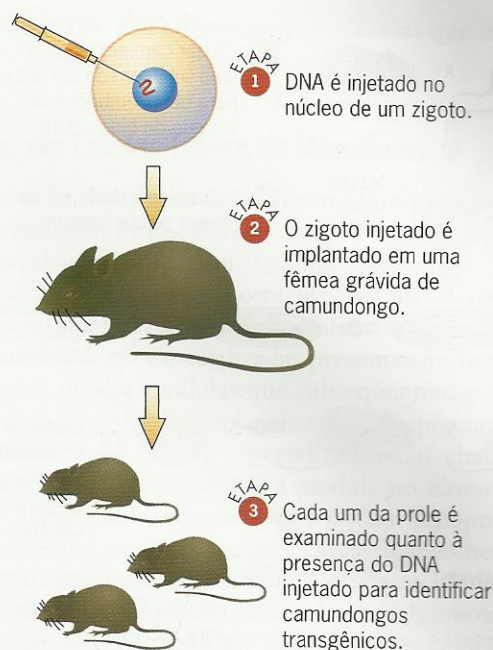
Como discutimos no Cap. 17, uma sequência de DNA “exógena” que foi inserida no genoma é referida como um transgene. Existem dois métodos gerais de introdução de transgenes em cromossomos de camundongo. Um é baseado na injeção de DNA em zigotos ou embriões (Fig. 22.27)





**Fig. 22.26** ■ Clonagem de um gene contendo uma mutação de inserção de um camundongo.

e o outro é baseado na injeção ou transfecção de DNA em grandes populações de células cultivadas que foram derivadas de embriões muito jovens de camundongos (Fig. 22.28). Tais **células tronco embrionárias** (ou **células ES**) vêm da massa celular interna, um grupo de células encontradas no estágio de blástula dos embriões de camundongo. Tais células podem ser cultivadas *in vitro*, transfectadas ou injetadas com DNA e então introduzidas em outros embriões de camundongo em desenvolvimento. Ao acaso, algumas das células ES introduzidas podem contribuir para a formação de tecidos adultos, de modo que, quando o camundongo nasce, ele pode consistir em uma mistura de dois tipos de



**Fig. 22.27** ■ Produção de camundongos transgênicos mediante injeção de ovos e implantação em fêmeas para completar seu desenvolvimento.

células, as suas próprias e as derivadas das células ES cultivadas (e potencialmente mutagenizadas). Tais camundongos são chamados de **quimeras**. Se as células ES contribuírem para a linhagem germinativa da quimera, uma mutação que foi introduzida nelas pela inserção de DNA exógeno tem uma chance de ser transmitida para a geração seguinte. O cruzamento de um camundongo quimérico pode, portanto, estabelecer uma linhagem transgênica mutante. Linhagens transgênicas também podem ser produzidas cruzando-se camundongos derivados de embriões que foram diretamente injetados com DNA. Hoje em dia, usando um destes métodos, milhares de linhagens transgênicas foram criadas, e, entre elas, muitas mutações significativas do ponto de vista do desenvolvimento foram detectadas.

Uma aplicação especial da mutagenese transgênica é possível com genes que já foram clonados. Para tais genes, é possível criar mutações que causam o nocaute do funcionamento gênico. Uma **mutação "nocaute"** pode ajudar um pesquisador a determinar que papel o gene normal desempenha durante o desenvolvimento. Para criar uma mutação nocaute, a sequência do gene clonado deve ser alterada *in vitro* e então introduzida em células ES cultivadas. Em baixa frequência, o transgene mutado substituirá seu alelo normal no cromossomo por recombinação homóloga, um processo chamado de **targeting**. Células ES que contêm uma mutação nocaute alvo podem ser usadas para criar quimeras, que podem então ser cruzadas para produzir linhagens transgênicas que levam a mutação nocaute. É então uma simples questão de cruzamento de dois heterozigotos para tornar a mutação nocaute homozigota e determinar que efeito ela tem





**Fig. 22.28** ■ Produção de camundongos transgênicos por tecnologia de células tronco embrionárias (ES).

no desenvolvimento. Como exemplo, consideremos o efeito de uma mutação nocaute no gene *Hoxc-8*. Os camundongos que são homozigotos para esta mutação desenvolvem um par extra de costelas posteriores ao conjunto normal de costelas. Eles também têm dedos fechados em suas patas anteriores. O fenótipo de costelas extras nestes camundongos mutantes é remanescente de transformações segmentais que são vistas com mutações homeóticas em *Drosophila*.

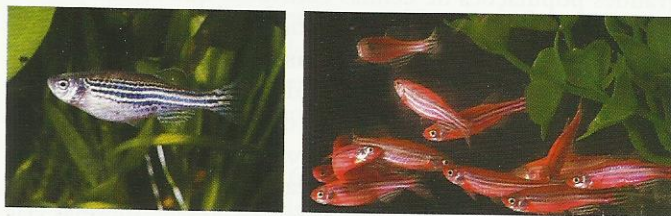
A análise genética do desenvolvimento em camundongos nos dá indícios do desenvolvimento em nossa própria espécie. Por exemplo, mutações em pelo menos dois genes diferentes de camundongos mimetizam o desenvolvimento de assimetrias anormais no eixo esquerdo-direito em seres humanos. Normalmente, seres humanos, camundongos e outros vertebrados apresentam estruturas que são assimétricas ao longo do eixo corpóreo esquerdo-direito. O tubo cardíaco sempre se inclina para a direita, e o fígado, o estômago e outras vísceras são mudados para a esquerda ou

para a direita da linha média corpórea. Em mutantes, essas assimetrias características não são vistas, talvez devido a um defeito nos mecanismos que estabelecem o plano corpóreo básico. O estudo destes tipos de mutantes no camundongo pode portanto ajudar a elucidar a maneira como os órgãos são posicionados em seres humanos.

### O Peixe-Zebra: Mutações Nocaute Morfolino

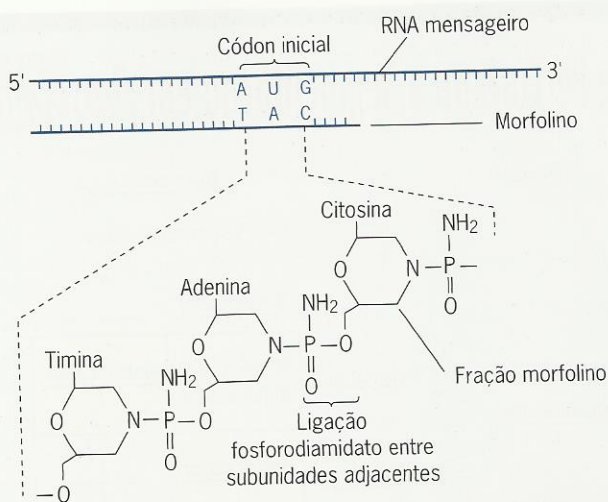
Nos anos 1960 e 1970, George Streisinger, um geneticista de bacteriófagos, voltou suas atenções para a análise genética de *Danio rerio*, um pequeno peixe de água doce nativo do sudeste da Ásia. Devido a suas listras horizontais, esta espécie é comumente conhecida como peixe-zebra (Fig. 22.29). Por muitas décadas, os peixes-zebra foram tidos como ornamentos em aquários. Eles são criados facilmente e, em condições ótimas, podem completar uma geração em apenas cinco a seis meses. Streisinger acreditou que o peixe-zebra podia ser usado como um organismo modelo para o estudo da fisiologia e do desenvolvimento de vertebrados. Sua crença foi concretizada nos anos 1990 quando pesquisadores em vários países começaram projetos para induzir mutações no peixe-zebra. Dentro de alguns anos, foram coletadas milhares de linhagens mutantes. O que torna o peixe-zebra especialmente atraente a geneticistas que estudam o desenvolvimento é que os óvulos do peixe são transparentes. Além disso, eles são fertilizados externamente. Assim, é possível observar os eventos de desenvolvimento embrionário sem ter de extrair os óvulos das fêmeas (como tem de ser feito com o camundongo) ou sem ter de dissecar embriões em desenvolvimento.

O genoma do peixe-zebra contém aproximadamente 1,5 bilhão de pares de bases de DNA – cerca de metade do genoma humano. O seqüenciamento desse DNA indica que o genoma do peixe-zebra contém cerca de 22.000 genes. As funções destes genes estão sendo analisadas por mutagenese do mesmo modo que geneticistas analisam genes em *Drosophila* e *C. elegans*. Entretanto, devido ao ciclo de vida mais longo do peixe-zebra, os resultados destas análises surgiram lentamente. Para acelerar o progresso, os geneticistas do peixe-zebra desenvolveram procedimentos para mimetizar os efeitos de mutações em genes específicos. Estes procedimentos não induzem mutações, isto é, não alteram o DNA em si. Eles bloqueiam ou atenuam a expressão da informação



**Fig. 22.29** ■ O peixe-zebra, *Danio rerio*. Peixe tipo selvagem à esquerda; peixe com um transgene que expressa a proteína vermelha fluorescente derivada de anêmonas-do-mar à direita.





**Fig. 22.30** ■ Morfolino com uma sequência de pares de bases ao redor do códon de início de um RNA mensageiro. Morfolinos tipicamente têm 20 a 25 subunidades de tamanho. Cada subunidade consiste em uma fração morfolino ligada a uma base nitrogenada (A, T, U, G ou C). As subunidades são ligadas umas às outras por ligações fosforodiamidato em vez de ligações fosfodiéster, como o são no DNA e no RNA.

codificada no DNA. Essa perturbação da expressão gênica é obtida injetando-se em embriões do peixe-zebra moléculas sintéticas que se ligam a sequências específicas em moléculas de RNA mensageiro e impedem que esses mRNA sejam traduzidos em polipeptídeos (Fig. 22.30). O material injetado é um DNA estável que é criado para fazer pares de bases especificamente com a região 5' não traduzida e o códon inicial de um mRNA. Como o arcabouço açúcar-fosfato deste DNA análogo é feito de frações de morfolina, a molécula injetada é chamada de **morfolino**. Menos de 10 nanogramas de um morfolino injetados no vitelo de um óvulo fertilizado de peixe-zebra ou embrião inicial (contendo menos de 16 células) são eficazes para perturbar a expressão gênica por um considerável período de tempo. Os animais que se desenvolvem destes ovos ou embriões injetados têm fenótipos como os de animais que têm mutações de fato no gene alvo. Assim, eles são **fenocópias** do fenótipo mutante. Com o término da sequência do genoma do peixe-zebra, os geneticistas podem agora usar os morfolinos para inativar a expressão gênica e avaliar se um determinado produto gênico tem um papel significativo no desenvolvimento.

A técnica morfolino tem sido usada para inativar a expressão gênica nos embriões de outros vertebrados, incluindo o sapo *Xenopus laevis*, o camundongo e as galinhas. É uma técnica anti-sentido que se desenvolveu em paralelo às técnicas de interferência no RNA discutidas no Cap. 21. Entretanto, ao contrário da interferência no RNA, os morfolinos não causam a degradação do RNA mensageiro. Eles funcionam inibindo o começo da tradução, provavelmente prejudicando a ligação dos ribossomos ao RNA mensageiro ao qual são complementares. Também, os morfolinos são moléculas

unifilamentares. Em contraste, as moléculas que induzem o RNA de interferência são bifilamentares.

## Estudos em Células Tronco de Mamíferos

As células terminalmente diferenciadas no corpo humano – linfócitos, neurônios, fibras musculares e assim por diante – em geral não se dividem. Quando estes tipos de células são perdidos por meio de morte, eles devem ser repostos ou o tecido ao qual pertencem irá atrofiar. A reposição ocorre quando células não especializadas presentes no tecido se dividem para produzir células que subsequentemente se diferenciam no tipo celular especializado. Tais precursores não especializados de células especializadas são chamados de **células tronco**. Por exemplo, a medula no fêmur humano contém células indiferenciadas que podem repopular vários tipos de células sanguíneas. Tais células tronco hematopoiéticas mantêm o sistema circulatório abastecido de linfócitos, eritrócitos e plaquetas. Os tecidos em alguns órgãos tais como o coração parecem ter muito poucas células tronco; conseqüentemente, sua capacidade de regenerar material perdido ou danificado é limitada. Outros tecidos, tais como o revestimento do tubo digestivo e a pele, têm grandes populações de células tronco que substituem vigorosamente células diferenciadas à medida que elas são perdidas. Como células tronco desses tipos são encontradas em organismos desenvolvidos, elas são chamadas de células tronco adultas.

Células tronco também são encontradas em organismos em desenvolvimento. De fato, durante os primeiros estágios do desenvolvimento, todas as células, ou a sua maioria, têm propriedades de células tronco. Nós já vimos, por exemplo, que células obtidas da massa celular interna de um embrião de camundongo podem ser cultivadas *in vitro* e subsequentemente transplantadas para outro embrião de camundongo, onde irão dividir-se e finalmente contribuir para a formação de muitos tipos de tecidos e órgãos. As células tronco embrionárias têm, portanto, um tremendo potencial de desenvolvimento; isto é, elas são **pluripotentes** – capazes de se desenvolver de muitos modos.

Independentemente de serem derivadas de tecido embrionário ou adulto, as células tronco criam uma oportunidade para estudar os mecanismos envolvidos na diferenciação de tipos celulares especiais. Células tronco podem ser obtidas de uma variedade de mamíferos, incluindo camundongos, macacos e humanos. Elas podem ser cultivadas *in vitro* e examinadas para diferenciação enquanto crescem lá ou após serem transplantadas para um organismo hospedeiro. Enquanto em cultura, as células tronco podem ser tratadas de vários modos para avaliar o que ativa seu desenvolvimento em uma direção específica. Técnicas moleculares, incluindo tecnologias de *chip* gênico, permitem que pesquisadores determinem quais genes as células estão expressando à medida que seus programas de desenvolvimento avançam.

Como as células tronco embrionárias (ES) têm o maior potencial de desenvolvimento, elas são idealmente adequadas a este tipo de análise. Tais células geralmente são deriva-





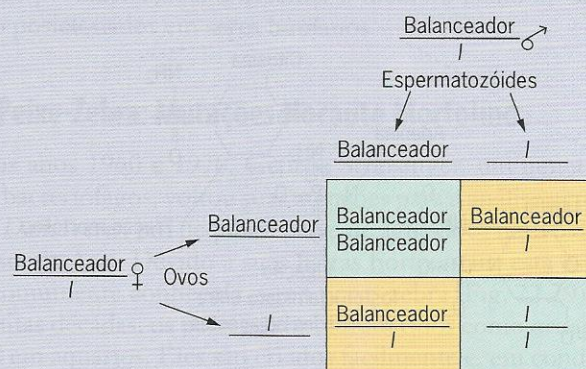
## UM MARCO NA GENÉTICA: Mutações que Perturbam a Segmentação em *Drosophila*

Os corpos de muitos tipos de animais são construídos a partir de segmentos. Por exemplo, os insetos possuem três partes corpóreas principais – uma cabeça, um tórax e um abdome –, e o tórax e o abdome são feitos de unidades segmentares menores. Mesmo os vertebrados exibem uma anatomia segmentada, especialmente durante os primeiros estágios do desenvolvimento quando massas discretas de tecidos chamadas somitos se formam em intervalos regulares ao longo do arcabouço. A segmentação parece portanto ser uma característica padrão da estrutura de muitos animais.

Anatomistas e embriologistas aprenderam sobre os modos pelos quais os segmentos se formam e se alinham dissecando animais em desenvolvimento. Os geneticistas também aprenderam sobre a segmentação fazendo dissecções. Entretanto, eles não usaram bisturis ou facas; em vez disso, empregaram mutações para definir e analisar os genes que estão envolvidos no processo de segmentação.

A dissecção genética da segmentação teve um grande impulso em 1980 quando Christiane Nüsslein-Volhard e Eric Wieschaus publicaram um trabalho descrevendo mutações que alteram o número, o tamanho e a composição dos segmentos em embriões e larvas de *Drosophila*.<sup>1</sup> Nüsslein-Volhard e Wieschaus sabiam que a organização espacial dos embriões de *Drosophila* é influenciada pelos produtos dos genes de efeito materno. Esses produtos gênicos são depositados nos ovócitos e, após a fertilização, eles orientam o desenvolvimento do embrião ao longo de seus eixos anterior-posterior e dorsal-ventral. Nüsslein-Volhard e Wieschaus também sabiam que os estágios posteriores do desenvolvimento de *Drosophila* envolvem os produtos de genes homeóticos tais como *Ultrabithorax*. Estes produtos gênicos especificam os destinos de segmentos individuais no corpo. Algumas regiões diferenciavam-se em segmentos torácicos e outras diferenciavam-se em segmentos abdominais. Nüsslein-Volhard e Wieschaus reconheceram que deve haver uma outra classe de produtos gênicos que atuam após os produtos gênicos fornecidos maternamente terem estabelecido o plano básico corpóreo, mas antes que os produtos dos genes homeóticos apareçam. Eles procuraram definir tais produtos gênicos triando mutações que perturbam a segmentação em embriões e larvas de *Drosophila*.

Uma vez que tais mutações atuariam como letais recessivos, Nüsslein-Volhard e Wieschaus criaram um esquema para induzi-las tratando *Drosophila* adulta com o mutágeno etilmetanossulfonato. Cromossomos mutagenizados foram recuperados na condição heterozigota na geração seguinte, e então, por uma série de cruzamentos envolvendo cromossomos balanceadores, eles foram tornados homozigotos para determinar se novas mutações letais recessivas haviam sido induzidas. Em um experimento no qual 5.800 cromossomos mutagenizados foram triados, 4.500 portavam novas mutações letais recessivas. Uma vez identificados, estes cromossomos mutantes foram mantidos na condição heterozigota em estoques balanceados.



**Fig. 1 ■** Segregação de embriões homozigotos para uma mutação letal recessiva recém-induzida (l) em um estoque balanceado de *Drosophila*. O cromossomo balanceador leva um marcador dominante e múltiplas inversões para suprimir recombinação entre ele e o cromossomo com a mutação letal. A maioria dos cromossomos balanceadores também levam uma mutação letal recessiva que não é alélica à mutação letal no outro cromossomo. Assim, a prole que é homozigota para o cromossomo balanceador bem como as que são homozigotas para a mutação letal recém-induzida morrem durante o desenvolvimento.

Os estoques letais balanceados segregam homozigotos letais recessivos a cada geração (Fig. 1). Se a mutação letal em um estoque atua no início do desenvolvimento, os embriões homozigotos para ela não eclodirão do ovo. Assim, na pesquisa por mutações que afetam o desenvolvimento inicial, Nüsslein-Volhard e Wieschaus examinaram embriões que não eclodiram em sua coleção de estoques. Eles procuraram intensamente embriões nos quais o padrão segmentar tinha sido perturbado.

Esta árdua pesquisa foi bem-sucedida. Nüsslein-Volhard e Wieschaus identificaram mutações perturbadoras de segmentação em 15 genes diferentes. Mutações em três dos genes criaram espaços segmentais nos embriões, mutações em seis dos genes deletaram partes alternadas do padrão segmentar e mutações em seis dos genes deletaram partes de segmentos e substituíram estas partes por duplicatas de outras partes (Fig. 2). Os genes *gap*, os genes *pair-rule* e os genes *segment polarity* foram então definidos. Trabalhos subsequentes mostraram que os produtos destes genes desempenham papéis centrais no estabelecimento do padrão segmentar em embriões de *Drosophila*. Homólogos dos produtos destes genes também funcionam no desenvolvimento embriológico de outros organismos.

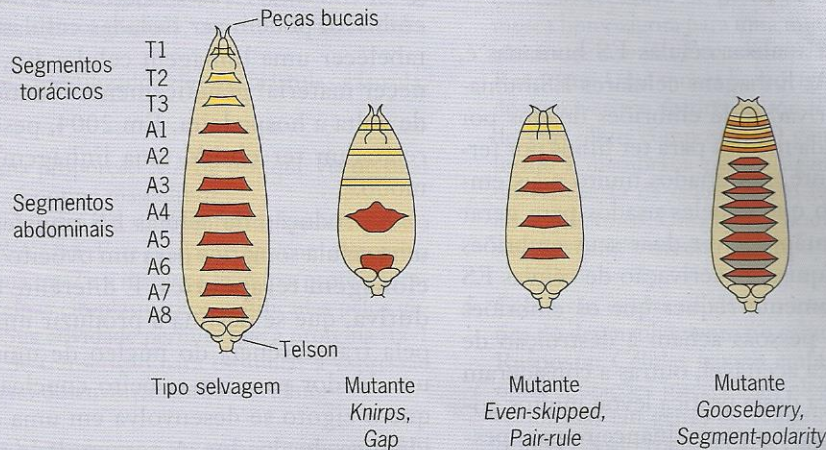
Por este trabalho que abriu caminhos, Nüsslein-Volhard e Wieschaus receberam o Prêmio Nobel de Fisiologia ou Medicina de 1995. Edward Lewis também dividiu este prêmio por seu trabalho pioneiro sobre genes homeóticos de *Drosophila*.

<sup>1</sup>Nüsslein-Volhard, C., e E. Wieschaus. 1980. Mutations affecting segment number and polarity in *Drosophila*. *Nature* 287: 795-801.





## UM MARCO NA GENÉTICA (continuação)



**Fig. 2 ■** Embriões de *Drosophila* mostrando os fenótipos de mutações em três genes de segmentação diferentes: *knirps* (um gene gap), *even-skipped* (um gene pair-rule) e *gooseberry* (um gene de polaridade segmentar). As áreas coloridas são bandas de denticulos (projeções curtas filamentosas da pele do embrião) associadas à parte anterior de cada segmento torácico e abdominal no embrião tipo selvagem. Essas bandas de denticulos são um meio para identificar segmentos e porções de segmentos em embriões em desenvolvimento. No mutante *knirps*, o padrão de denticulo indica que grande parte do material em segmentos abdominais adjacentes está faltando. No mutante *even-skipped*, ele mostra que ao longo do corpo faltam segmentos alternados e, no mutante *gooseberry*, ele mostra que a parte posterior de cada segmento foi deletada e substituída por uma cópia em imagem especular da porção anterior, que é marcada pelos denticulos. As cópias em imagem especular da porção anterior de cada segmento são mostradas em laranja no tórax e em cinza no abdome. Fonte: Nüsslein-Volhard & Wieschaus, 1980. *Nature* 287: 795-801. Figura 1, p. 796.

### QUESTÕES PARA DISCUSSÕES

1. Nüsslein-Volhard e Wieschaus procuravam fenótipos mutantes em embriões não eclodidos. Antes que eles tivessem iniciado seu trabalho, a maioria dos geneticistas de *Drosophila* procuraram fenótipos mutantes em adultos. O que isto nos diz sobre a importância do “ponto de vista” ou “plano de visão” na procura de novos conhecimentos?
2. Uma vez que os genes de segmentação foram definidos por mutações, eles tinham de ser caracterizados em nível molecular por clonagem, análise de expressão e assim em diante. Uma vez caracterizados em nível molecular, sondas feitas a partir destes genes puderam ser usadas para estudar genes homólogos em outros organismos. Que aspectos envolvendo os genes de segmentação seriam interessantes para estudar outros organismos segmentados tais como crustáceos e platelmintos e organismos não segmentados tais como água-viva e estrela-do-mar?

das da massa celular interna de embriões que foram criados por fertilização *in vitro*. As células isoladas desta massa são plaqueadas em uma camada de “células alimentadoras” mitoticamente inativas, que fornecem fatores de crescimento para estimular a divisão. Para células ES de camundongo que crescem em cultura, o tempo de duplicação é de cerca de 12 horas; para células ES humanas, ele é de cerca de 36 horas. Após as células embrionárias isoladas terem crescido por um tempo enquanto em cultura nas células alimentadoras, elas são dissociadas e replaqueadas para estabelecer populações de células tronco clonais, que podem então ser congeladas para estocagem por longo prazo. Uma população de células clonais é a que surgiu de uma única célula geradora.

As células ES começam a se diferenciar quando são transferidas de culturas de células alimentadoras para suspensões de culturas com um meio apropriado. Nessas condições, elas formam **corpos embrionários**, que são agregados multicelulares constituídos de células diferenciadas e indiferenciadas. Para algumas espécies, os corpos embrionários assemelham-se a embriões pré-blastocisto. As células desses corpos podem diferenciar-se nos tipos de células especializadas que são derivadas de cada uma das três camadas germinativas primárias – ectoderma, mesoderma e endoderma. Por exemplo, elas podem formar neurônios, que são derivados do ectoderma; células de músculos lisos ou células cardíacas de contração rítmica, que são derivadas do mesoderma; ou células



das ilhotas pancreáticas, que são derivadas do endoderma. Observando este processo em linhagens celulares diferentes – por exemplo, em linhagens nas quais determinados genes foram mutados –, pode ser possível dissecar a rede genética de interações envolvida na diferenciação de vários tipos de células.

A questão de procurar e analisar células ES humanas é, logicamente, controversa. As linhagens celulares ES humanas hoje em uso foram derivadas de embriões doados por pessoas que procuraram ajuda médica para ter filhos por fertilização *in vitro*. Tipicamente, são criados muito mais embriões por este processo do que aqueles usados para gerar crianças. Um casal pode então decidir doar seus embriões não usados para fins de pesquisa. A derivação de células ES de tais embriões necessariamente requer que os embriões sejam destruídos. Algumas pessoas acham a destruição de embriões iniciais uma prática aceitável; outras a consideram imoral. A controvérsia acerca desta prática levou muitos governos a suspender ou restringir suporte financeiro para pesquisas em células tronco embrionárias humanas. Nos EUA, por exemplo, o apoio do governo federal só é dado a projetos que usam linhagens de células tronco humanas estabelecidas antes de 9 de agosto de 2001. Os fundos para projetos que usam linhagens celulares estabelecidas após esta data devem vir de outras fontes.

O debate sobre financiamento para pesquisa de células tronco embrionárias humanas foi intensificado pela perspectiva de usar células ES humanas para curar doenças que resultam da perda de tipos celulares específicos, tais como diabetes mellitus (em que as células das ilhotas pancreáticas foram perdidas) e doença de Parkinson (na qual alguns tipos de neurônios em determinada região do cérebro foram perdidos). A terapia com células ES também foi proposta para tratar incapacidades tais como as resultantes de danos na coluna dorsal. A idéia é transplantar células derivadas de células ES para o tecido morto ou danificado e deixar que tais células regenerem as partes perdidas ou danificadas do tecido. Experimentos com camundongos e ratos sugerem que esta estratégia pode funcionar em humanos. Entretanto, muitos problemas técnicos ainda não foram resolvidos. Por exemplo, ainda não é possível obter culturas puras de um determinado tipo celular diferenciado. Quando células ES humanas se desenvolvem em cultura, elas se diferenciam em muitos tipos de células. Isolar um tipo – digamos, por exemplo, células cardíacas – é um incrível desafio técnico.

Os proponentes da terapia com células tronco humanas também têm que resolver outros tipos de problemas. As células derivadas de uma cultura *in vitro* podem dividir-se descontroladamente e formar tumores ao serem transplantadas para um hospedeiro ou podem ser eliminadas pelo sistema imunológico do hospedeiro. Para evitar este último problema, os pesquisadores propuseram transplantar células que são geneticamente idênticas às células do hospedeiro. Tais células geneticamente idênticas poderiam ser criadas usando-se uma das células somáticas

do hospedeiro para gerar a população de células ES. Uma célula somática do hospedeiro poderia ser fusionada com um ovócito enucleado obtido de uma mulher doadora (não necessariamente a hospedeira). Se o zigoto alterado geneticamente, que é diplóide, divide-se para formar um embrião, podem ser isoladas células do embrião para estabelecer uma linhagem celular ES, que pode então fornecer material geneticamente idêntico para o transplante de volta à hospedeira. Em 2004, pesquisadores na Coreia relataram ter obtido uma linhagem de células ES deste modo.

A produção de células ES transferindo-se o núcleo de uma célula somática para um ovócito enucleado é chamada **clonagem terapêutica**. É diferente da **clonagem reprodutiva**, que se destina a produzir um indivíduo completo pela transferência do núcleo de uma célula somática de um doador para um ovócito enucleado e então deixando que o zigoto se desenvolva em uma cópia geneticamente idêntica do doador. A comunidade científica internacional não considera a clonagem reprodutiva de humanos segura. A principal preocupação é que alguns dos genes no núcleo da célula somática transferido possam ter sido modificados por metilação ou algum aspecto da organização da cromatina – considere, por exemplo, os genes em um cromossomo X inativo. Estas modificações serão alteradas após o núcleo da célula somática entrar no ovócito enucleado? Os genes transferidos serão reprogramados para orientar o desenvolvimento apropriado do embrião? Experimentos com outros mamíferos – ovelhas, camundongos e gatos – sugerem que os embriões produzidos por transferência de núcleo somático nem sempre se desenvolvem normalmente. Por conseguinte, há uma ampla concordância de que a clonagem reprodutiva de humanos não deve ser tentada.

#### ■ PONTOS IMPORTANTES

- Muitos genes de vertebrados, como os genes *Hox*, foram identificados por homologia com genes isolados de um organismo modelo tal como *Drosophila* e *C. elegans*.
- Entre os vertebrados, o camundongo e o peixe-zebra fornecem oportunidades de estudar mutações que afetam o desenvolvimento.
- Camundongos transgênicos podem ser criados injetando-se DNA em ovócitos ou embriões ou inserindo células tronco embrionárias transfectadas em embriões em desenvolvimento.
- Camundongos transgênicos podem ser a fonte de mutações insercionais, incluindo mutações nocaute, em genes que têm um significado do ponto de vista do desenvolvimento.
- A injeção de morfólino anti-sentido em ovócitos ou embriões de vertebrados pode perturbar a expressão gênica e produzir fenocópias mutantes que revelam o significado desenvolvimental do gene alvo.
- Células tronco de mamíferos, especialmente as derivadas de embriões, podem ser cultivadas *in vitro* para estudar os mecanismos subjacentes à diferenciação.



## Exercícios Básicos Ilustram a análise genética básica.

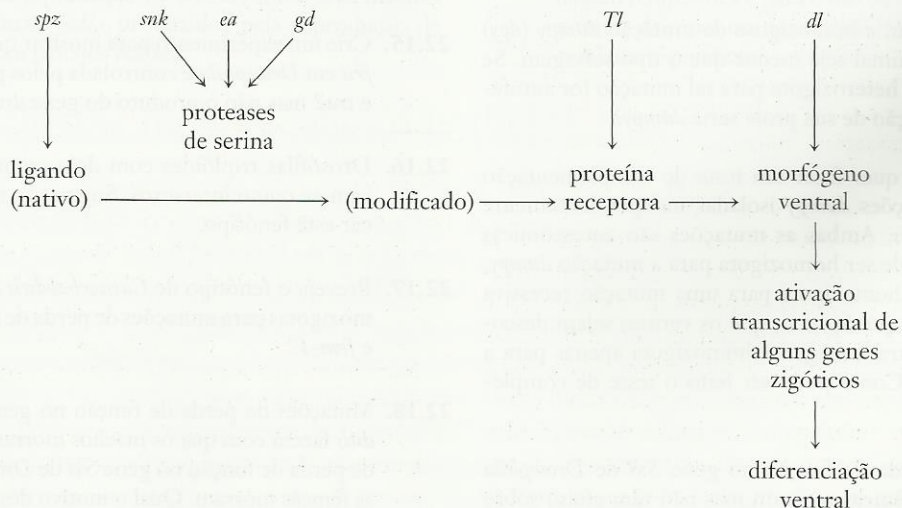
1. Arrume os seguintes estágios do desenvolvimento de *Drosophila melanogaster* em ordem cronológica desde os primeiros até os últimos: pupa, blástula, zigoto, gástrula, ovócito não fertilizado, larva, adulto.  
**Respostas:** ovócito não fertilizado, zigoto, blástula, gástrula, larva, pupa, adulto.
2. Em *C. elegans* com uma proporção X:A de 0,5, que via de sinal de transdução causa o fenótipo masculino?  
**Resposta:** A proteína her-1 é um sinal secretado que é expresso quando a proporção X:A é 0,5. Este sinal interage com a proteína receptora de membrana tra-2 e a inativa. Quando a proteína tra-2 é inativada, as proteínas citoplasmáticas codificadas pelos genes *fem* são ativadas. Sua atividade permite que a proteína tra-1 atue como um fator de transcrição para estimular a expressão dos genes envolvidos no desenvolvimento masculino.
3. Fêmeas de *Drosophila* homozigotas para uma mutação autossômica recessiva recém-descoberta põem ovos que não eclodem em larvas, independente do genótipo de seus machos. Entretanto, as próprias fêmeas não apresentam anomalias óbvias. Que tipo de gene esta mutação define?  
**Resposta:** A nova mutação define um gene de efeito materno.
4. Preveja o fenótipo ocular de uma mosca homozigota para uma mutação recessiva de perda de função no gene *sevenless*. Uma mosca homozigota para uma mutação recessiva de perda de função no gene *bride of sevenless* teria o mesmo fenótipo?  
**Resposta:** Uma mosca homozigota para a mutação *sevenless* não desenvolveria o fotorreceptor R7 em cada um dos omatídeos de seus olhos compostos. O gene *sevenless* codifica o receptor ligado à membrana para o ligando extracelular que dispara a diferenciação da célula R7; o ligando é codificado pelo gene *bride of sevenless*. Uma mosca homozigota para a mutação *bride of sevenless* teria o mesmo fenótipo.
5. Como são usadas células tronco embrionárias na análise genética do desenvolvimento no camundongo?  
**Resposta:** Células tronco embrionárias (ES) são usadas para obter mutações insercionais aleatórias e mutações nocaute direcionadas em genes de camundongo. Em ambos os casos, as células ES são transfectadas com DNA mutado enquanto mantidas em cultura. As células transfectadas são então fusionadas com embriões jovens de camundongo para criar quimeras, que são implantadas em mães adotivas e deixadas desenvolver-se em adultos. A criação apropriada dos adultos permite que sejam identificados os efeitos das mutações recém-induzidas.

## Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. O produto protéico do gene *dorsal* (*dl*) em *Drosophila* foi denominado morfógeno ventral, isto é, uma substância que causa a formação de estruturas ventrais no embrião em virtude de sua alta concentração nos núcleos do lado ventral do blastoderma. Entretanto, a proteína dorsal pode entrar nestes núcleos ventrais apenas se um receptor na superfície ventral do embrião tiver sido ativado. Este receptor é codificado pelo gene *Toll* (*Tl*). O ligando extracelular do receptor Toll é codificado pelo gene *spätzle* (*spz*). Este ligando pode existir em dois estados, "nativo" e "modificado", e o estado modificado é necessário

para a ativação do receptor Toll. Os produtos de três genes, *snake* (*snk*), *easter* (*ea*) e *gastrulation defective* (*gd*), são necessários para converter o ligando nativo no ligando modificado. Todos os três produtos gênicos são proteases de serina, proteínas capazes de cortar outras proteínas em algumas serinas na cadeia polipeptídica. Usando estes fatos, faça um diagrama da via de desenvolvimento que faz com que a proteína dorsal induza a formação de estruturas ventrais no embrião de *Drosophila*.

**Resposta:** Aqui está uma representação.





O produto protéico do gene *spz* é modificado pelas proteases de serina feitas pelos genes *snk*, *ea* e *gd*. Em sua forma modificada, este ligando é capaz de ativar a proteína receptora Toll, mas a ativação é restrita ao lado ventral do embrião. Quando o receptor Toll foi ativado (supostamente unindo-se ao ligando spätzle modificado), ele transduz um sinal para o citoplasma do embrião. Este sinal finalmente faz com que a proteína dorsal se mova para os núcleos do lado ventral do embrião, onde atua como um fator de transcrição para regular a expressão dos genes zigóticos envolvidos na diferenciação dos destinos ventrais.

2. Considerando a via descrita acima, quais seriam os fenótipos das mutações recessivas de perda de função nos genes *spz* e *Tl*?

**Resposta:** Para fins de referência, devemos notar que as mutações de perda de função em *dl* são letais de efeito materno, isto

é, embriões de mães *dl/dl* morrem durante o desenvolvimento. Quando estes embriões que morrem são examinados, eles não apresentam estruturas ventrais. Os geneticistas dizem que eles são “dorsalizados”. Este fenótipo peculiar é devido à falha do fator de transcrição dorsal em induzir o desenvolvimento apropriado nos núcleos ventrais do embrião. Na ausência desta indução, as células ventrais diferenciam-se como se estivessem do lado dorsal do embrião. Espera-se que mutações em *spz* e *Tl* tenham o mesmo efeito fenotípico porque bloqueariam etapas na via que acaba por fazer com que a proteína dorsal induza diferenciação ventral. Mutações recessivas em *spz* e *Tl* são portanto letais de efeito materno. Fêmeas homozigotas para estas mutações produzem embriões dorsalizados que morrem durante o desenvolvimento.

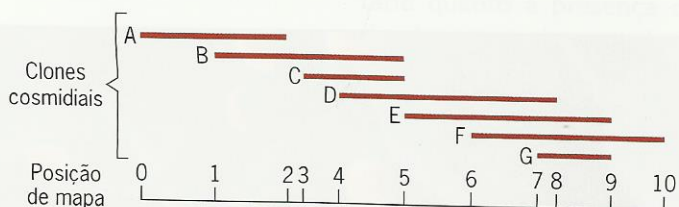
## Questões e Problemas

**Acentuam a compreensão e desenvolvem as habilidades analíticas.**

- 22.1. Que eventos formam a blástula durante o desenvolvimento animal? Que eventos formam a gástrula?
- 22.2. Durante a ovocitogênese, que mecanismos enriquecem o citoplasma de ovócitos animais com materiais nutritivos e determinativos?
- 22.3. Faça distinção entre “determinação” e “diferenciação”.
- 22.4. As “camadas germinativas primárias” são componentes da linhagem germinativa? Explique.
- 22.5. Esquematize as principais etapas na análise genética do desenvolvimento em um organismo modelo tal como *Drosophila* ou *Caenorhabditis*.
- 22.6. Por que o embrião inicial de *Drosophila* é um sincício?
- 22.7. Em *Drosophila*, que tecidos larvais produzem os órgãos externos do adulto?
- 22.8. Em *Caenorhabditis*, a homozigose da mutação *dumpy* (*dpy*) faz com que o animal seja menor que o tipo selvagem. Se um hermafrodita heterozigoto para tal mutação for autofecundado, que fração de sua prole seria *dumpy*?
- 22.9. Um pesquisador quer fazer um teste de complementação entre duas mutações *dumpy* isoladas independentemente em *Caenorhabditis*. Ambas as mutações são autossômicas recessivas. Além de ser homozigota para a mutação *dumpy*, uma linhagem é homozigota para uma mutação recessiva ligada ao X (*unc*) que faz com que os vermes sejam descoordenados. A outra linhagem é homozigota apenas para a mutação *dumpy*. Como pode ser feito o teste de complementação?
- 22.10. Mutações de perda de função no gene *Sxl* de *Drosophila* fazem com que fêmeas morram mas não têm efeito sobre machos. Que fenótipo uma mutação de ganho de função mais provavelmente teria?
- 22.11. Um pesquisador de *Drosophila* descobriu uma mutação autossômica recessiva que faz com que animais XX se desenvolvam em machos estéreis. Proponha um esquema para descobrir se a mutação é um alelo do gene *tra* ou do gene *tra2*.
- 22.12. Faça um diagrama da via de determinação do sexo em *Drosophila*.
- 22.13. Que fenótipo você esperaria em uma drosófila XX que é homozigota para mutações de perda de função nos genes *tra* e *dsx*? Explique.
- 22.14. Preveja o fenótipo sexual (anatômico e comportamental) dos seguintes genótipos mutantes em *Drosophila*: (a) XY; *tra fru/tra fru*; (b) XX; *tra fru/tra fru*; (c) XY; *dsx/dsx; fru/fru*; (d) XX; *dsx/dsx; fru/fru*.
- 22.15. Crie um experimento para mostrar que a expressão do gene *fru* em *Drosophila* é controlada pelos produtos dos genes *tra* e *tra2* mas não o produto do gene *dsx*.
- 22.16. Drosófilas triplóides com dois cromossomos X desenvolvem-se como intersexos. Sugira um mecanismo para explicar este fenótipo.
- 22.17. Preveja o fenótipo de *Caenorhabditis* XO e XX que são homozigotas para mutações de perda de função nos genes *tra-1* e *fem-1*.
- 22.18. Mutações de perda de função no gene *xol-1* de *Caenorhabditis* fazem com que os machos morram, enquanto mutações de perda de função no gene *Sxl* de *Drosophila* fazem com que as fêmeas morram. Qual o motivo destes efeitos opostos?



- 22.19. Uma pesquisadora está tentando clonar o gene *dpy-3* de *Caenorhabditis*. Mutações recessivas neste gene fazem com que os vermes sejam menores que o normal, um fenótipo referido como “dumpy”. A pesquisadora mapeou o gene *dpy-3* em relação a outros genes em um dos cromossomos de *Caenorhabditis* e, usando os dados do mapa, obteve clones de cosmídeos com DNA tipo selvagem da vizinhança do gene *dpy-3*. Para determinar quais clones contêm o gene *dpy-3*, a pesquisadora injetou cada um deles nas gônadas de hermafroditas homozigotos para uma mutação recessiva em *dpy-3*. Ela examinou a prole transgênica desses hermafroditas injetados quanto ao fenótipo *dumpy*. Pelos resultados mostrados adiante, onde no mapa molecular a pesquisadora deve colocar o gene *dpy-3*?



| Clone Cosmicial | Fenótipo de Prole Transgênica |
|-----------------|-------------------------------|
| A               | dumpy                         |
| B               | dumpy                         |
| C               | dumpy                         |
| D               | tipo selvagem                 |
| E               | tipo selvagem                 |
| F               | dumpy                         |
| G               | dumpy                         |

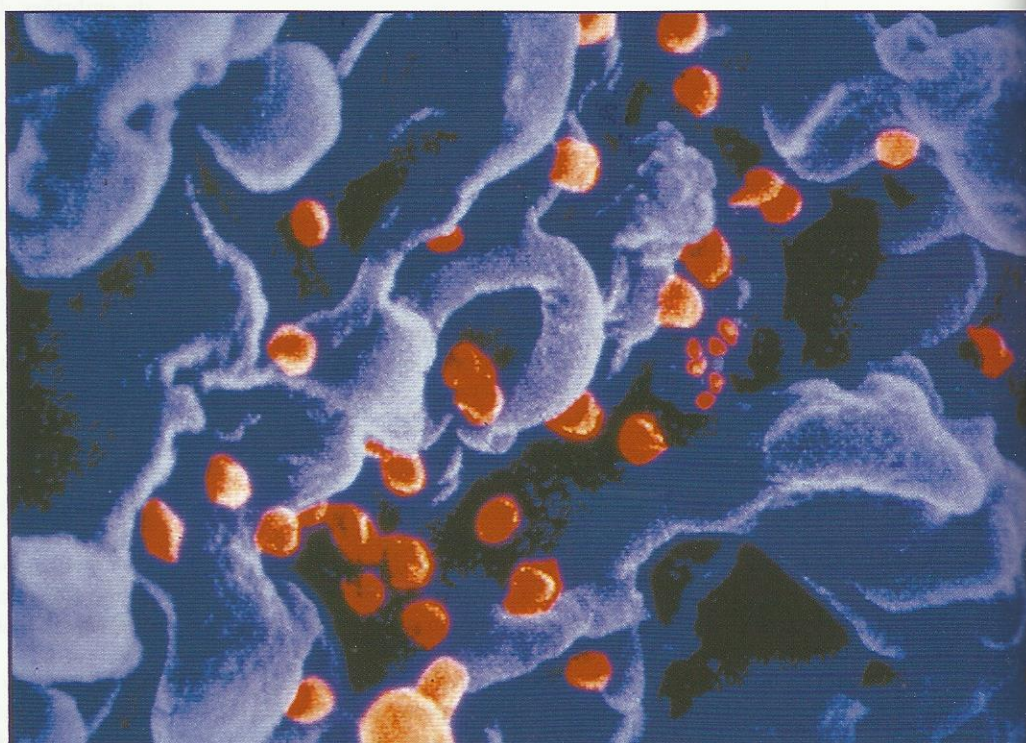
- 22.20. Como *dorsal*, *bicoid* é um gene estritamente de efeito materno em *Drosophila*; isto é, não tem expressão zigótica. Mutações recessivas em *bicoid* (*bcd*) causam morte embrionária por impedir a formação de estruturas anteriores. Preveja os fenótipos de (a) animais *bcd/bcd* produzidos pela reprodução de machos e fêmeas heterozigotos; (b) animais *bcd/bcd* produzidos pela reprodução de fêmeas *bcd/bcd* com machos *bcd/+*; (c) animais *bcd/+* produzidos pela reprodução de fêmeas *bcd/bcd* com machos *bcd/+*; (d) animais *bcd/bcd* produzidos pela reprodução de fêmeas *bcd/+* com machos *bcd/bcd*; (e) animais *bcd/+* produzidos pela reprodução de fêmeas *bcd/+* com machos *bcd/bcd*.
- 22.21. Em *Drosophila*, mutações recessivas no gene *dorsal* (*dl*) do eixo dorsal-ventral causam um fenótipo dorsalizado em embriões produzidos por mães *dl/dl*; isto é, não desenvolvem-se estruturas ventrais. Preveja o fenótipo de embriões produzidos por fêmeas homozigotas para uma mutação recessiva no gene *nanos* do eixo anterior-posterior.
- 22.22. Um pesquisador está planejando coletar mutações em genes de efeito materno que controlam os primeiros eventos no desenvolvimento de *Drosophila*. Que fenótipo o pesquisador deve procurar nesta pesquisa de mutações de efeito materno?
- 22.23. Um pesquisador está planejando coletar mutações nos genes *gap*, que controlam as primeiras etapas na segmentação de embriões de *Drosophila*. Que fenótipo o pesquisador deve procurar em sua pesquisa de mutações no gene *gap*?
- 22.24. Por que mulheres, mas não homens, que são homozigotas para o alelo mutante que causa fenilcetonúria têm filhos que são física e mentalmente retardados?
- 22.25. Faça um diagrama de uma via que mostre as contribuições dos genes *sevenless* (*sev*) e *bride of sevenless* (*boss*) para a diferenciação do fotorreceptor R7 nos omatídeos dos olhos de *Drosophila*. Onde *eyeless* (*ey*) se encaixaria nesta via?
- 22.26. Quando o gene *Pax6* de camundongo, que é homólogo ao gene *eyeless* de *Drosophila*, é expresso em *Drosophila*, ele produz olhos compostos extras com omatídeos, como os olhos normais de *Drosophila*. Se o gene *eyeless* de *Drosophila* fosse introduzido em camundongos e expresso lá, que efeito você esperaria? Explique.
- 22.27. Como você pode mostrar que dois genes *Hox* de camundongo são expressos em tecidos diferentes e em momentos diferentes durante o desenvolvimento?
- 22.28. A metilação do DNA, a acetilação de histonas e a compactação do DNA na cromatina por certos tipos de proteínas às vezes são referidas como modificações epigenéticas do DNA. Estas modificações criam dificuldades para clonagem reprodutiva. Elas também criam dificuldades para clonagem terapêutica e para o uso de células tronco destinadas a tratar doenças ou danos que envolvem a perda de tipos celulares específicos?



# O Controle Genético do Sistema Imunológico de Vertebrados

## Tópicos

- A Resposta Imune: Uma Visão Geral
- Componentes do Sistema Imunológico de Mamíferos
- A Resposta Imune Mediada por Anticorpos (Humoral)
- A Resposta Imune Mediada por Células T (Celular)
- Memória Imunológica
- Rearranjos Genômicos Durante Diferenciação do Linfócito B
- Mudança de Classe de Anticorpo
- Montagem de Genes de Receptores de Células T por Recombinação Somática
- Regulação da Expressão de Genes de Imunoglobulina



Vírus da imunodeficiência humana (vermelho) infectando um linfócito T4 (azul).

## A AIDS Enfatizou o Papel Vital do Sistema Imunológico

Em 1971, quando Brandon ainda era bem pequeno, ele caiu no degrau em frente a sua casa e cortou o joelho. Como o sangramento não parava, sua mãe o levou para a emergência do hospital local. Testes sanguíneos revelaram que Brandon tinha hemofilia A, uma deficiência hereditária do fator VIII de coagulação sanguínea (Cap. 5). Pessoas com este distúrbio recessivo ligado ao X exibem sangramento excessivo mesmo por pequenos cortes. Entretanto, o sangramento pode ser interrompido por tratamento com o fator de coagulação ausente. Em 1982, quando Brandon se machucou em um acidente de automóvel, ele recebeu transfusões de sangue e tratamento com fator VIII de coagulação. Ele se recuperou das contusões e continuou tendo uma vida



ativa normal. Oito anos depois, Brandon começou a ter fadiga, febre recorrente e perda de peso. Ele parecia contaminar-se por qualquer vírus com que tivesse contato. Logo depois, foi feito o temido diagnóstico: Brandon tinha síndrome de deficiência imunológica adquirida (AIDS), a doença fatal causada pelo vírus da imunodeficiência humana (HIV) (Cap. 18). Alguns meses depois, Brandon contraiu um tipo de pneumonia causada por *Pneumocystis carinii*. Este protozoário é um patógeno oportunista que infecta pessoas – tais como pacientes com AIDS – cujos sistemas imunológicos estão enfraquecidos. O HIV destrói o sistema imunológico de seus hospedeiros matando um tipo de glóbulos brancos que são necessários para uma resposta imunológica normal. Brandon morreu menos de um ano depois.

Brandon foi infectado pelo HIV no sangue contaminado usado na transfusão após o acidente de automóvel. Antes de 1985, o soro usado nas transfusões não era testado quanto à presença do HIV, que foi identificado em 1983. Hoje, todo sangue doado é testado para detectar HIV antes de ser depositado nos bancos de sangue. Assim, o risco de contrair AIDS por uma transfusão de sangue é extremamente baixo – cerca de um caso em  $10^5$  transfusões.

A AIDS destaca o papel vital que o sistema imunológico desempenha em nossa sobrevivência. Muitos de nós presenciaremos a saúde de um parente, amigo ou vizinho com AIDS deteriorar-se ante nossos olhos após o HIV destruir células essenciais de seu sistema imunológico. Caso não, observamos os horríveis efeitos da AIDS sobre o corpo humano por meio das notícias na mídia. A dor e o sofrimento vivenciados pelos pacientes terminais de AIDS ilustram o que ocorre quando o sistema imunológico não pode mais desempenhar o seu papel. Como resultado, a maioria de nós conhece os papéis importantes que nosso sistema imunológico exerce em nossas vidas. Diariamente nosso sistema imunológico luta contra uma ampla gama de microrganismos patogênicos.

O ambiente em que vivemos contém um enorme número de vírus, bactérias, fungos e outros agentes diferentes. Estes organismos, muitos dos quais são infecciosos, estão constantemente evoluindo e produzindo novas linhagens. Os vírus, em particular, acumulam novas mutações e evoluem muito rapidamente (Cap. 18). Como, então, o sistema imunológico pode proteger-nos de todos esses patógenos diferentes, alguns dos quais são novos e confrontam-se com células do sistema imunológico pela primeira vez? Neste capítulo, responderemos a esta pergunta. Examinaremos os componentes do sistema imunológico humano e os processos pelos quais as respostas imunológicas nos protegem de agentes infecciosos. Ao fazer isto, encontraremos alguns genes fascinantes que são montados a partir de pedaços de genes durante a diferenciação das células que os contêm. Veremos como este processo de montagem facilita a produção de uma variedade de proteínas com a especificidade necessária para combater uma aparente ilimitada variedade de agentes infecciosos. Por motivo de abreviação e foco, restringiremos nossa discussão aos sistemas imunológicos de mamíferos, com ênfase especial nos seres humanos.

## A RESPOSTA IMUNE: UMA VISÃO GERAL

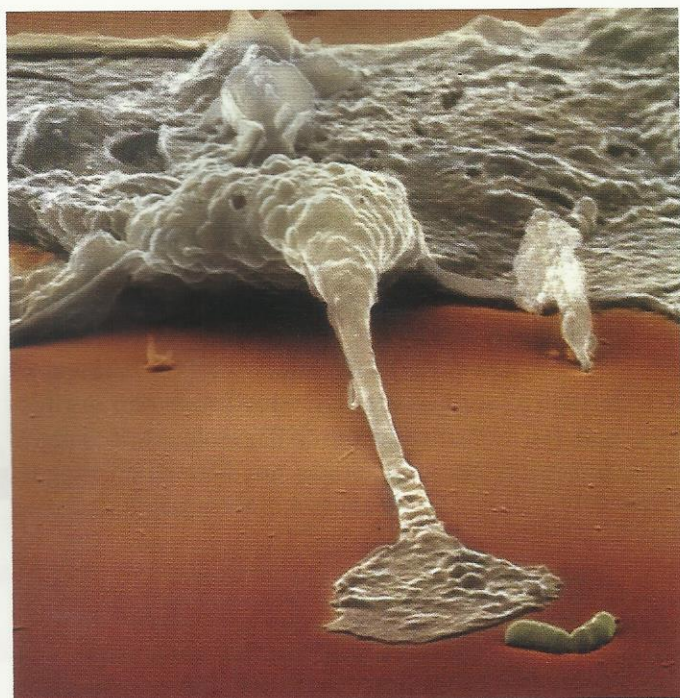
A resposta imune em mamíferos envolve três etapas: (1) reconhecimento da substância exógena, (2) comunicação desse reconhecimento às células que respondem e (3) eliminação do agente invasor pelas atividades coordenadas de proteínas especiais e células especializadas.

Quando um mamífero é infectado por um vírus, uma bactéria, um fungo ou outras células invasoras patogênicas, o corpo monta uma complexa resposta imune. Cada resposta imune envolve três etapas essenciais:

1. *Reconhecimento* da entidade exógena.
2. *Comunicação* do reconhecimento às células apropriadas.
3. *Eliminação* da entidade invasora.

Alguns componentes da resposta imune são dirigidos contra agentes infecciosos em geral e constituem as **respostas imunes inespecíficas**. Elas incluem inflamação dos tecidos infectados que resultam em aumento do fluxo sanguíneo e recrutamento de um tipo especial de células sanguíneas chamadas *fagócitos* (Fig. 23.1) que ingerem e destroem vírus, bactérias e fungos de modo inespecífico. Entretanto, as respostas mais importantes são específicas para a substância exógena e são portanto denominadas **respostas imunes específicas**. No restante deste capítulo, enfocaremos as duas principais respostas imunes específicas: a síntese de proteínas especiais e a produção de células que protegem os mamíferos de agentes infecciosos. Devemos destacar que estas duas respostas não ocorrem independentemente. Em vez disso, as células envolvidas nas duas respostas devem comunicar-se umas com as outras para que as respostas sejam eficazes.





**Fig. 23.1** ■ Uma projeção de fagócito captando uma bactéria em bastonete.

De fato, as respostas imunológicas a agentes exógenos são processos complexos, e devemos estar familiarizados com as células e proteínas que estão envolvidas antes de discutir estes processos em mais detalhes.

### PONTOS IMPORTANTES

- A resposta imune em mamíferos envolve três etapas: (1) reconhecimento da substância exógena, (2) comunicação desse reconhecimento às células de resposta e (3) eliminação do agente invasor.
- Respostas inespecíficas incluem o recrutamento de fagócitos para ingerir e destruir vírus e microrganismos.
- Respostas específicas incluem a síntese de proteínas especiais e células que trabalham juntas para remover a entidade exógena de nosso corpo.

## COMPONENTES DO SISTEMA IMUNOLÓGICO DE MAMÍFEROS

A resposta imune em mamíferos envolve as atividades coordenadas de vários tipos de glóbulos brancos, alguns produzindo proteínas de ligação a antígenos chamadas anticorpos e outros desenvolvendo-se em células que destroem células apresentadoras de antígenos.

A característica mais marcante do sistema imunológico é sua **especificidade**. Essa especificidade é proporcionada

por (1) um grupo de células especializadas, cada uma fazendo seu papel de modo altamente coordenado; (2) dois exércitos de proteínas, *anticorpos* e *receptores de células T*, cada um com capacidade ilimitada de reconhecer substâncias exógenas; e (3) um conjunto especializado de proteínas, os *antígenos de histocompatibilidade*, que facilitam a comunicação entre células e permitem que células do sistema imunológico reconheçam substâncias e distingam suas próprias células de células exógenas (não-próprias). As células e proteínas mais importantes deste sistema imunológico são citadas no Quadro 23.1. No total, o sistema imunológico de um humano adulto consiste em cerca de um trilhão ( $10^{12}$ ) de glóbulos brancos e  $10^{15}$  ou mais moléculas proteicas especiais, todas agindo em conjunto para eliminar substâncias exógenas.

Quando um agente exógeno tal como um vírus ou bactéria invade o corpo humano, ele é reconhecido por células do sistema imune e ativa a produção de anticorpos e receptores de células T que reconhecem especificamente a substância exógena e ligam-se a ela. Uma substância que é ligada a um anticorpo ou receptor de célula T é chamada **antígeno** (substância geradora de anticorpo), e um antígeno que provoca uma resposta imune é denominado **imunógeno**. Praticamente todas as macromoléculas biológicas, incluindo polipeptídeos, polissacarídeos, ácidos nucleicos e muitas substâncias sintéticas, podem agir como antígenos, e a maioria pode funcionar como imunógenos.

A resposta imune de mamíferos tem dois componentes principais: a resposta mediada por anticorpo, ou resposta humoral (com base na palavra em latim que significa “líquido” – donde líquidos corpóreos), e a resposta mediada por células T, ou resposta celular. A resposta imune humoral envolve a produção e a secreção de anticorpos no sistema circulatório, onde tais anticorpos se ligam a seus respectivos antígenos e os sequestram. Os complexos anticorpo/antígeno são então ingeridos e destruídos por uma classe especial de glóbulos brancos. O sistema imune humoral é primariamente responsável por proteger mamíferos de células exógenas (por exemplo, bactérias e fungos) e de vírus livres antes que eles infectem células do organismo hospedeiro. A resposta imune celular envolve a produção de receptores de células T que revestem as superfícies dos linfócitos T, permitindo que esses glóbulos brancos especiais reconheçam e destruam as células infectadas do organismo hospedeiro. Assim, o sistema imune celular é amplamente responsável por proteger os mamíferos de infecções virais. As células e macromoléculas que participam das respostas imunes de mamíferos são descritas nas seções seguintes.

### Células Especializadas Medeiam a Resposta Imune

A resposta imune a uma substância exógena é baseada nas atividades coordenadas de uma variedade de células altamente especializadas. Todas essas células são derivadas de células precursoras indiferenciadas, chamadas **células tronco**, presentes na medula óssea. Um componente importante da resposta imune são os **linfócitos B**, denominados **células B**



## QUADRO 23.1

## Componentes do Sistema Imunológico de Mamíferos

| Componente                                                    | Descrição e Função                                                                                                                                                                                                    |
|---------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>I. Proteínas</b>                                           |                                                                                                                                                                                                                       |
| Antígenos                                                     | Substâncias que disparam uma resposta imune; substâncias geradoras de anticorpos                                                                                                                                      |
| Anticorpos                                                    | Proteínas produzidas pelo sistema imunológico que se ligam a antígenos e ajudam na sua destruição                                                                                                                     |
| Receptores de células T                                       | Proteínas produzidas pelo sistema imunológico em resposta a antígenos; elas se tornam ancoradas à superfície de células T <i>killer</i> e ligam antígenos com a ajuda de antígenos principais de histocompatibilidade |
| Antígenos do complexo principal de histocompatibilidade (MHC) | Proteínas de superfície celular que (1) permitem que células do sistema imunológico distingam substâncias exógenas do que é próprio e (2) facilitam a comunicação entre células                                       |
| <b>II. Células</b>                                            |                                                                                                                                                                                                                       |
| Células tronco                                                | Células indiferenciadas da medula óssea que dão origem a várias células do sistema imunológico                                                                                                                        |
| Fagócitos                                                     | Células grandes que capturam, ingerem e destroem agentes exógenos invasores tais como vírus, bactérias e fungos                                                                                                       |
| Macrófagos                                                    | Células fagocíticas que ingerem antígenos e os apresentam na superfície da célula para interações com outras células do sistema imunológico                                                                           |
| Linfócitos B                                                  | Células que se diferenciam (na medula óssea de mamíferos) em plasmócitos produtores de anticorpos e células B de memória                                                                                              |
| Plasmócitos                                                   | Leucócitos produtores de anticorpos derivados de linfócitos B                                                                                                                                                         |
| Células B de memória                                          | Células B que facilitam a produção rápida de um determinado anticorpo em resposta a um segundo e subsequente encontro com um antígeno                                                                                 |
| Linfócitos T                                                  | Células que sofrem eventos importantes de diferenciação no timo e depois se diferenciam nos vários tipos de células T                                                                                                 |
| Células T <i>helper</i>                                       | Células T que respondem à apresentação de um antígeno por um macrófago, estimulando linfócitos B a produzir anticorpos e linfócitos T a produzir receptores de células T                                              |
| Células T supressoras                                         | Células T que suprimem a produção de anticorpos e receptores de células T por células B e T, respectivamente                                                                                                          |
| Células T citotóxicas ( <i>killer</i> )                       | Células T que levam receptores de células T e destroem células que apresentam os antígenos reconhecidos                                                                                                               |
| Células T de memória                                          | Células T que facilitam a produção rápida de um determinado receptor de célula T no segundo e subsequentes encontros com um antígeno                                                                                  |

porque, em aves, elas maturam na Bursa de Fabricius, uma estrutura do tipo saco que desempenha algumas das mesmas funções da medula óssea em mamíferos. O outro componente principal da resposta imune é amplamente desempenhado pelos **linfócitos T**, denominados **células T** porque sofrem importantes eventos de diferenciação na glândula timo. As raízes gregas da palavra *linfócito* significam célula incolor, um nome apropriado para essas pequenas células brancas sanguíneas. O fenótipo quase incolor desses linfócitos contrasta marcadamente com o fenótipo altamente pigmentado de seus vizinhos contendo hemoglobina, as hemácias. Um outro grupo de células do sistema imunológico são os **fagócitos** (do grego, que significa “comer células”). Um tipo de fagócito, o **macrófago** (do grego, que significa “grande comedor”), tem um papel central em iniciar ambos os componentes da resposta imune.

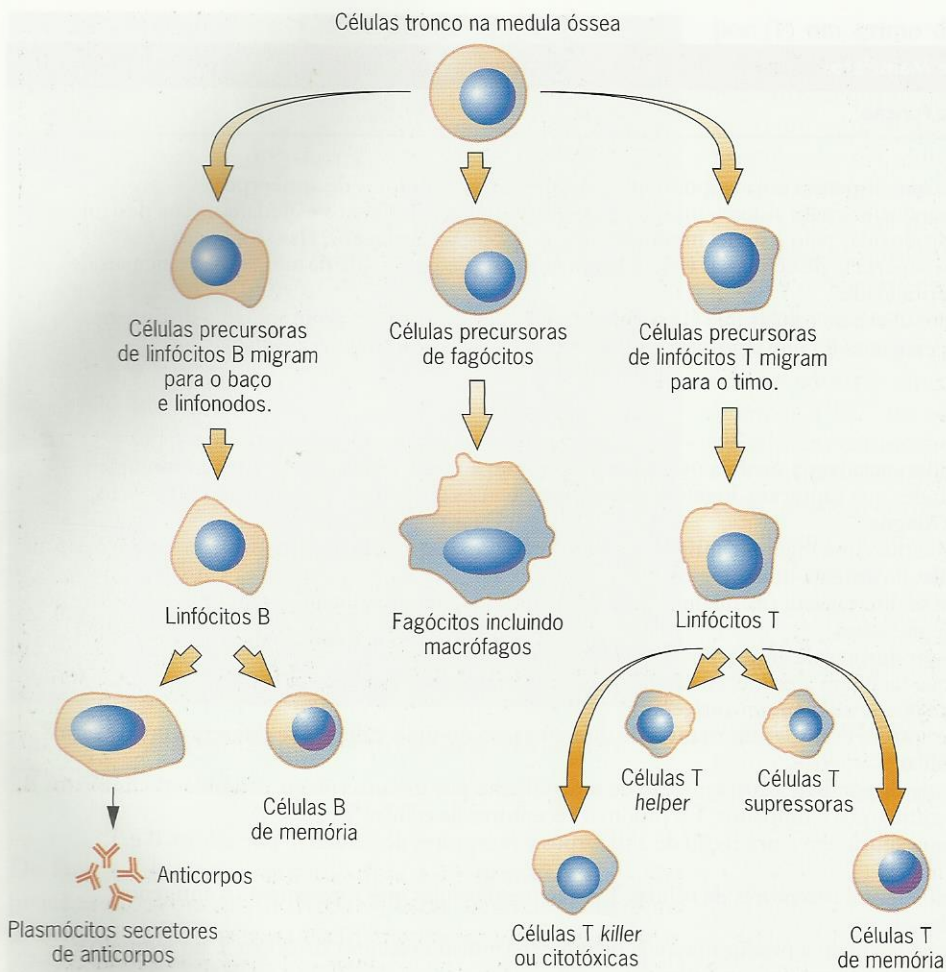
Ao receber os sinais apropriados, os linfócitos B diferenciam-se em **plasmócitos**, que produzem proteínas de ligação a antígenos chamadas anticorpos, e **células B de memória**, que facilitam uma produção mais rápida de anticorpos durante uma exposição posterior ao mesmo antígeno. Os linfócitos T diferenciam-se em quatro tipos importantes de células T.

1. **Células T *killer* ou citotóxicas** levam proteínas de ligação a antígeno chamadas *receptores de células T* em suas superfícies celulares e, como o nome indica, as células *killer* apresentam os antígenos apropriados.
2. **Células T *helper*** estimulam linfócitos B e T a se diferenciar em plasmócitos produtores de anticorpos e células T *killer*, respectivamente.
3. **Células T supressoras** ajudam a “suprimir” ou atenuar a atividade de plasmócitos e células T *killer*.
4. **Células T de memória** “lembram” do antígeno e proporcionam uma produção rápida de células T *killer* durante encontros subsequentes com o mesmo antígeno.

A Fig. 23.2 resume as origens e funções das células envolvidas em aspectos importantes da resposta imune em mamíferos.

Algumas pessoas herdaram ou adquiriram defeitos nas células do sistema imune. Os mais extremos destes distúrbios são a **síndrome da imunodeficiência grave combinada (SCID; herdada)** e a **síndrome da imunodeficiência adquirida (AIDS; discutida no Cap. 18)**. Pessoas com SCID são incapazes de montar uma resposta imune de qualquer tipo,





**Fig. 23.2** ■ As células mais importantes do sistema imunológico são derivadas de células tronco da medula óssea por três linhagens celulares separadas. Elas dão origem (1) a células B e plasmócitos produtores de anticorpos, (2) aos vários tipos de células T e (3) aos fagócitos, incluindo macrófagos.

seja mediada por anticorpo ou mediada por células T. Essas pessoas não têm linfócitos B ou T funcionais. Existem várias formas de SCID, com cerca de metade dos casos causados por mutações recessivas ligadas ao X e metade por defeitos autossômicos recessivos.

### Proteínas Especializadas Fornecem Especificidade Imunológica

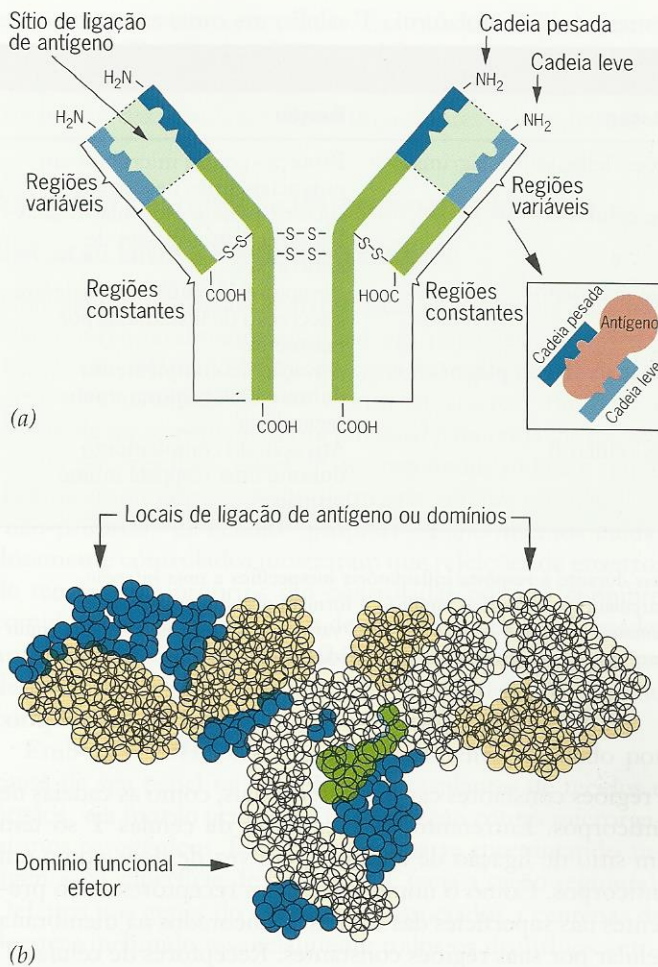
Especificidade imunológica é fornecida por duas classes de proteínas, anticorpos e receptores de células T, que medeiam os dois componentes principais – a resposta mediada por anticorpo ou humoral e a resposta mediada por células T ou celular – do sistema imunológico de mamíferos, respectivamente. Cada uma destas classes de proteínas apresenta uma variedade aparentemente ilimitada de sítios de ligação de antígeno, e os mecanismos pelos quais essa variabilidade é gerada são fascinantes.

Os **anticorpos** pertencem à classe de proteínas chamadas **imunoglobulinas**. Cada anticorpo é um tetrâmero composto de quatro polipeptídeos, duas **cadeias leves** idênticas e duas **cadeias pesadas** idênticas, unidas por

pontes dissulfídicas (Fig. 23.3a). As cadeias leves têm cerca de 220 aminoácidos de tamanho e as cadeias pesadas têm cerca de 440 a 450 aminoácidos de tamanho. Cada cadeia, pesada e leve, tem uma **região variável** amino-terminal, na qual a sequência de aminoácidos varia entre anticorpos específicos para antígenos diferentes, e uma **região constante** carboxila-terminal, na qual a sequência de aminoácidos é a mesma para todos os anticorpos de uma determinada classe de imunoglobulina (Ig), independente da especificidade de ligação do antígeno. As regiões variáveis de todas as cadeias de anticorpos têm cerca de 110 aminoácidos de tamanho.

As regiões de proteínas que realizam funções particulares são chamadas de **domínios**. Cada anticorpo tem dois **sítios de ligação de antígeno** ou domínios, cada um dos quais é formado por regiões variáveis de uma cadeia leve e uma cadeia pesada (Fig. 23.3b). Esses domínios de regiões variáveis formam sítios de ligação que interagem com o antígeno de modo chave-e-fechadura (Fig. 23.4). Além disso, as regiões constantes das duas cadeias leves interagem para formar um terceiro domínio, chamado de **domínio de função efetora**, que é responsável pela interação do anticorpo com outros componentes do sistema imunológico.

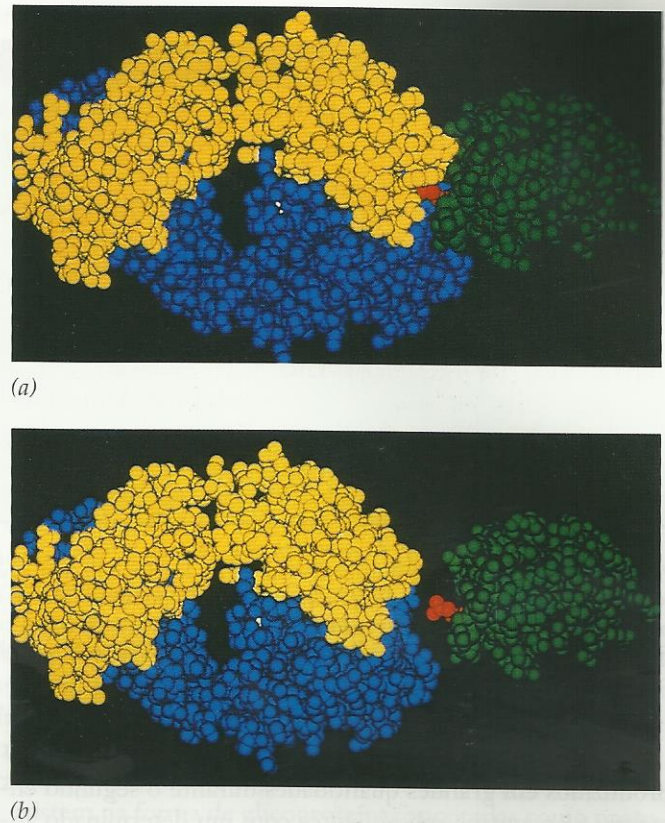




**Fig. 23.3** ■ (a) Diagrama e (b) modelo preenchido de estrutura do anticorpo. Cada anticorpo é um tetrâmero composto de quatro cadeias polipeptídicas: duas cadeias leves idênticas e duas cadeias pesadas idênticas. (a) Cada cadeia consiste em uma região variável (azul) e uma região constante (verde) e cada anticorpo tem dois sítios de ligação de antígeno, formados por regiões variáveis de cadeias leves e pesadas. O destaque mostra a interação “chave-e-fechadura” entre um antígeno e um anticorpo. (b) As duas cadeias pesadas são mostradas em branco e azul, respectivamente. As duas cadeias leves são mostradas em amarelo. Os componentes de cor verde são as frações carboidrato associadas; sua localização e estrutura irá variar dependendo da classe de imunoglobulina (Ig). A estrutura mostrada é para a molécula de IgG humana.

As cadeias leves dos anticorpos são de dois tipos, **cadeias kapa** e **cadeias lambda**, sendo o tipo determinado pela estrutura da região constante. Como veremos, os anticorpos podem ter a mesma especificidade de ligação de antígeno, como determinado pelas regiões variáveis das quatro cadeias, mas funções imunológicas diferentes, como determinado pelas regiões constantes das duas cadeias pesadas.

Existem cinco classes de anticorpos: **IgA**, **IgD**, **IgE**, **IgG**, e **IgM** (Quadro 23.2). A classe à qual pertence um anticorpo, e portanto a função que ele desempenha, é determinada pela estrutura de sua região constante da cadeia pesada (isto é, a



**Fig. 23.4** ■ Estrutura tridimensional de um complexo antígeno-anticorpo. É mostrado apenas um dos dois sítios de ligação de antígeno de um anticorpo típico. O antígeno (verde) é a enzima lisozima. O sítio de ligação de antígeno do anticorpo é formado pelas partes amino-terminais de uma cadeia leve (amarela) e uma cadeia pesada (azul). Uma glutamina que se projeta da lisozima no sítio onde o anticorpo se liga é mostrada em azul. As estruturas mostradas são baseadas em dados de difração de raios X. (a) O complexo total antígeno-anticorpo. (b) O antígeno e o anticorpo foram separados para mostrar suas estruturas complementares chave-e-fechadura.

estrutura de seu domínio de função efetora). Os anticorpos **IgA**, **IgD**, **IgE**, **IgG** e **IgM** têm cadeias pesadas  $\alpha$ ,  $\delta$ ,  $\epsilon$ ,  $\gamma$  e  $\mu$ , respectivamente. Em humanos, existem duas subclasses de anticorpos **IgA** e quatro subclasses de anticorpos **IgG**; cada subclasse tem cadeias pesadas diferentes mas proximamente relacionadas. A primeira classe de anticorpos produzidos em uma célula B em desenvolvimento é sempre **IgM**. Os anticorpos **IgM** iniciais são ancorados na membrana da célula, dando-lhe receptores de antígenos de superfície. Em uma *resposta imune primária* (o primeiro encontro com um determinado antígeno), uma célula B subsequentemente começará a produzir **anticorpos IgM secretados**, que são moléculas pentaméricas com dez sítios de ligação de antígenos. Depois, a célula B pode mudar para a produção de outras classes de anticorpos, um processo chamado de **mudança de classe**.

Os anticorpos **IgG** correspondem a cerca de 80% das moléculas de imunoglobulina no sangue humano. Eles são



## QUADRO 23.2

| Classes de Imunoglobulinas |               |                        |           |                                    |                                                                            |
|----------------------------|---------------|------------------------|-----------|------------------------------------|----------------------------------------------------------------------------|
| Tipo                       | Cadeia Pesada | Estrutura <sup>a</sup> | Proporção | Localização                        | Função                                                                     |
| IgA                        | $\alpha$      | $\alpha_2L_2$          | 14%       | Secreções: leite, saliva, lágrimas | Proteção contra micróbios em potenciais sítios de entrada                  |
| IgD                        | $\delta$      | $\delta_2L_2$          | 1%        | Sangue; células B                  | Incerta; podem estimular células B a fazer outras classes de anticorpos    |
| IgE                        | $\epsilon$    | $\epsilon_2L_2$        | <1%       | Tecidos; mastócitos <sup>b</sup>   | Receptores de antígeno que levam à secreção de histaminas por mastócitos   |
| IgG                        | $\gamma$      | $\gamma_2L_2$          | 80%       | Sangue; macrófagos, plasmócitos    | Ativação do complemento durante uma resposta imune secundária <sup>c</sup> |
| IgM                        | $\mu$         | $\mu_2L_2$             | 5%        | Sangue; células B                  | Ativação do complemento durante uma resposta imune primária <sup>c</sup>   |

<sup>a</sup>L representa uma cadeia leve de anticorpo, seja kapa ou lambda.

<sup>b</sup>Mastócitos são amplamente distribuídos em vários tecidos e liberam histaminas durante a resposta inflamatória inespecífica a uma infecção.

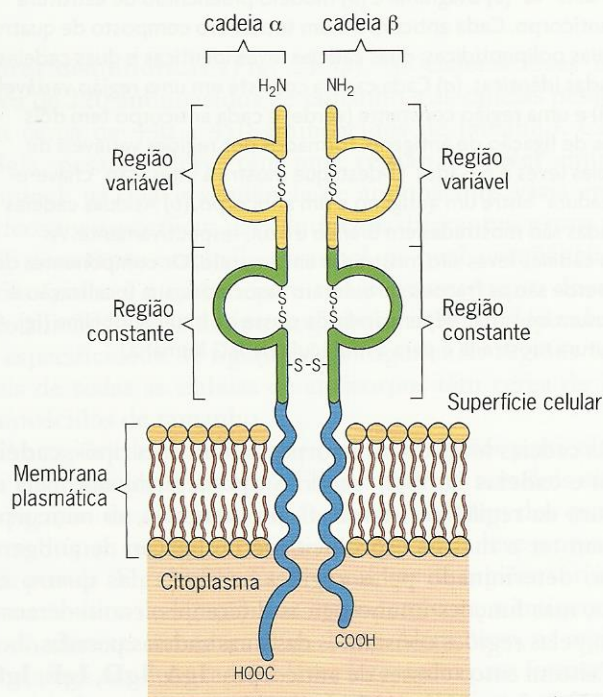
<sup>c</sup>Complemento é um sistema de aproximadamente 20 proteínas solúveis que circulam na corrente sanguínea e formam grandes complexos de proteína em resposta a complexos anticorpo/antígeno. As proteínas do complemento destroem células por uma variedade de mecanismos, tais como furos na membrana celular. Este sistema de proteínas é chamado "complemento" porque complementa a atividade de anticorpos em combater células exógenas patogênicas.

produzidos em grandes quantidades durante o segundo encontro e encontros subseqüentes com um determinado antígeno, a *resposta imune secundária*. Moléculas de IgA são as imunoglobulinas predominantes em secreções tais como saliva, lágrimas e leite, enquanto moléculas de IgE são encontradas como receptores de antígenos ligados a membrana em células especiais chamadas mastócitos em vários tecidos. A ligação de antígenos por estes receptores de superfície de mastócitos ativa a secreção de histaminas e outras amins, que são amplamente responsáveis pelas reações alérgicas observadas em pessoas com asma e febre do feno.

Se formos examinar a estrutura dos anticorpos, veremos que sua diversidade está quase que totalmente nas regiões variáveis das moléculas. Se tais polipeptídeos fossem sintetizados a partir de seqüências colineares de pares de nucleotídeos de genes, um gene por cadeia polipeptídica, o genoma teria de conter uma enorme gama de genes com seqüências altamente variáveis em uma ponta e seqüências essencialmente idênticas na outra ponta. Entretanto, este não é o caso. Os genes de anticorpos são estocados como segmentos de genes em células da linhagem germinativa, e tais segmentos gênicos são recompostos em genes durante a diferenciação de plasmócitos produtores de anticorpos. Examinaremos este processo nas seções subseqüentes deste capítulo.

A resposta imune celular é altamente específica do antígeno, como a resposta imune mediada por anticorpo, e é igualmente importante em defender mamíferos de infecções. Na resposta imune mediada por células, a especificidade do antígeno é fornecida por **receptores de células T**. Cada receptor de célula T é um dímero composto de um polipeptídeo  $\alpha$  e um polipeptídeo  $\beta$  (Fig. 23.5). Ambos os polipeptídeos têm regiões variáveis amino-terminais

e regiões constantes carboxila-terminais, como as cadeias de anticorpos. Entretanto, os receptores de células T só têm um sítio de ligação de antígeno, em vez de dois, como em anticorpos. Como o nome indica, tais receptores estão presentes nas superfícies das células T, ancorados na membrana celular por suas regiões constantes. Receptores de células T



**Fig. 23.5** ■ Estrutura de um receptor de célula T ancorado na membrana celular.



estão presentes tanto em células T citotóxicas (*killer*) quanto em células T *helper*, onde, em colaboração com os antígenos principais de histocompatibilidade, eles ligam antígenos e participam das atividades dessas importantes células imunes.

## Antígenos Principais de Histocompatibilidade: Distinção Entre o Próprio e o Exógeno

Transplantes de órgãos e enxertos de pele entre pessoas não aparentadas geralmente resultam em rejeições imunes dos tecidos transplantados. Em contraste, transplantes de órgãos entre gêmeos idênticos geralmente são aceitos. Também, a pele pode ser transplantada de um local para outro na mesma pessoa sem rejeição imune. Estes resultados indicam que o sistema imunológico pode distinguir células exógenas ou “não-próprias” de células “próprias”. Experimentos cuidadosamente controlados mostraram que rejeições de enxertos de tecidos não-próprios são controladas por um conjunto complexo de macromoléculas de superfície celular chamadas **antígenos de histocompatibilidade**. O mais importante destes antígenos de histocompatibilidade são as proteínas do **complexo principal de histocompatibilidade (MHC)**.

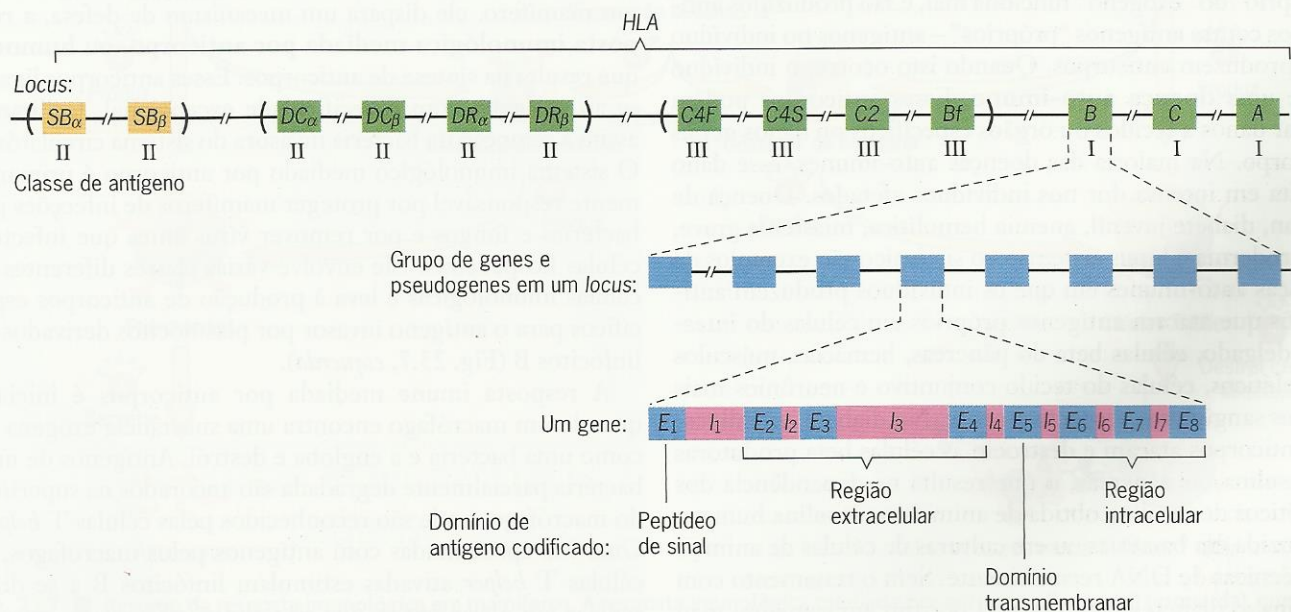
Embora o *MHC* tenha sido inicialmente estudado por causa de seu papel na rejeição de transplantes de tecidos e órgãos, sua função primária é dar proteção contra microrganismos patogênicos. Ele faz isto em parte apresentando antígenos na superfície de linfócitos T (veja a seção seguinte). Defeitos nos genes do *MHC* estão associados a centenas de doenças incluindo essencialmente todos os distúrbios auto-

imunes (veja adiante). Assim, o *locus MHC* estava entre as primeiras regiões do genoma humano a ser seqüenciada.

Muitas das proteínas codificadas pelo MHC são depositadas em superfícies celulares onde são altamente antigênicas para outros mamíferos bem como indivíduos “não idênticos” da mesma espécie. Em humanos, as proteínas do MHC são chamadas **antígenos HLA** (antígenos associados a leucócitos humanos [do inglês, *human-leukocyte-associated*]) porque foram descobertos nas superfícies dos leucócitos (glóbulos brancos). Os antígenos HLA são codificados por um grande grupo de genes, o **locus HLA**, no cromossomo 6 (Fig. 23.6). O grupo de genes *HLA* tem mais de  $3,5 \times 10^6$  pares de nucleotídeos de tamanho.

Os genes do *locus HLA* são altamente polimórficos, alguns tendo 100 ou mais alelos diferentes. Assim, a probabilidade de que dois indivíduos tenham os mesmos alelos de todos os genes de *HLA* é muito baixa. Como resultado, achar doadores de órgãos com antígenos HLA que correspondam exatamente aos do receptor é muito difícil.

Os genes de *HLA* codificam três classes diferentes de proteínas do MHC, cada classe de moléculas envolvida em um aspecto diferente da resposta imune. Os **genes classe I de HLA** codificam proteínas que são chamadas **antígenos de transplante** devido a seu papel na rejeição de tecidos durante transplantes de órgãos e tecidos. As proteínas da classe I existem na forma de glicoproteínas ancoradas como proteínas integrais da membrana com determinantes antigênicos expostos no lado externo das células. Elas estão presentes em praticamente todas as células de um organismo e fornecem às células T *killer* um mecanismo para distinguir o “exó-



**Fig. 23.6** ■ Organização do complexo principal de histocompatibilidade (*HLA*) no cromossomo humano 6. As posições relativas dos *loci* mapeados dentro deste enorme complexo gênico são mostradas em cima. A classe de antígeno de histocompatibilidade codificada por genes em cada *locus* é indicada logo abaixo do mapa. Cada *locus* dentro do complexo *HLA* contém vários genes e pseudogenes (centro). A estrutura de um típico gene classe I é mostrada embaixo. Note que éxons diferentes codificam domínios funcionais diferentes do produto gênico polipeptídico. Chave: E = éxon; I = intron.



geno” do “próprio”. Cada proteína da classe I do MHC pode ligar-se a um amplo espectro de antígenos e apresentá-los às células *T killer*.

Os **genes classe II de HLA** codificam polipeptídeos que estão situados primariamente nas superfícies de linfócitos B e macrófagos, células que interagem com células *T helper*. As proteínas classe II do MHC contêm sulcos de ligação de antígeno que se ligam a uma ampla gama de peptídeos correlatos e os apresentam a células *T helper*. Finalmente, os **genes classe III de HLA** codificam **proteínas do complemento** que interagem com complexos anticorpo/antígeno e ajudam a destruí-los por proteólise. As proteínas do complemento também destroem células infectadas perturbando a integridade de membranas celulares.

Com o término do seqüenciamento do cromossomo humano 6 em 2003, foi descoberto que alguns genes de *HLA* e outros de *MHC* ocorrem fora do *locus* clássico de *MHC* (*HLA*). Como resultado, o *locus* humano de *MHC* foi ampliado para cobrir  $7,6 \times 10^6$  pares de nucleotídeos. O *MHC* ampliado contém 421 genes e pseudogenes; 252 destas seqüências são classificadas como genes expressos. Além dos genes *HLA* clássicos discutidos anteriormente, o *MHC* ampliado contém grupos de genes codificantes de proteínas olfatórias (percepção de odores), histonas, proteínas de choque de calor, receptores de feromônios, fatores de necrose tumoral e outras proteínas e tRNA. O mapa completo dos genes no *locus* ampliado de *MHC* humano está disponível no número de dezembro de 2004 da *Nature Reviews* (página 891) e *online* em <http://www.nature.com/nrg/journal/v5/n12/poster/MHCmap>.

Às vezes, os mecanismos pelos quais o corpo distingue o “próprio” do “exógeno” funciona mal, e são produzidos anticorpos contra antígenos “próprios” – antígenos no indivíduo que produzem anticorpos. Quando isto ocorre, o indivíduo sofre uma **doença auto-imune**. Esses anticorpos podem causar danos a tecidos ou órgãos específicos ou danos gerais ao corpo. Na maioria das doenças auto-imunes, esse dano resulta em intensa dor nos indivíduos afetados. Doença de Crohn, diabetes juvenil, anemia hemolítica, miastenia grave, escleroderma e lúpus eritematoso sistêmico são exemplos de doenças auto-imunes em que os indivíduos produzem anticorpos que atacam antígenos próprios em células do intestino delgado, células beta do pâncreas, hemácias, músculos esqueléticos, células do tecido conjuntivo e neurônios mais células sanguíneas, respectivamente. No diabetes juvenil, esses anticorpos atacam e destroem as células beta produtoras de insulina no pâncreas, o que resulta na dependência dos diabéticos de insulina obtida de animais ou insulina humana produzida em bactérias ou em culturas de células de animais por técnicas de DNA recombinante. Sem o tratamento com insulina, o diabetes juvenil resulta por fim em cegueira, mau funcionamento renal, doença cardíaca e morte prematura.

#### ■ PONTOS IMPORTANTES

- A resposta imune de mamíferos envolve as atividades coordenadas de vários leucócitos especializados.

- Após exposição a uma substância exógena (antígeno), linfócitos B diferenciam-se em plasmócitos que produzem proteínas de ligação de antígenos chamadas anticorpos.
- Similarmente, linfócitos T desenvolvem-se em células *T killer* que levam receptores de antígenos de células T em suas superfícies e destroem as células apresentando os antígenos reconhecidos.
- Antígenos principais de histocompatibilidade altamente polimórficos nas células permitem que células *T killer* distingam células exógenas de células “próprias”, as células do indivíduo que montam a resposta imune.
- Indivíduos com doenças auto-imunes produzem anticorpos contra antígenos “próprios”.

## A RESPOSTA IMUNE MEDIADA POR ANTICORPOS (HUMORAL)

Os macrófagos degradam parcialmente vírus, bactérias e outras entidades exógenas invasores e apresentam os antígenos resultantes em suas superfícies, onde são reconhecidos pelas células *T helper*, que, por sua vez, estimulam linfócitos B a se diferenciar em plasmócitos produtores de anticorpos.

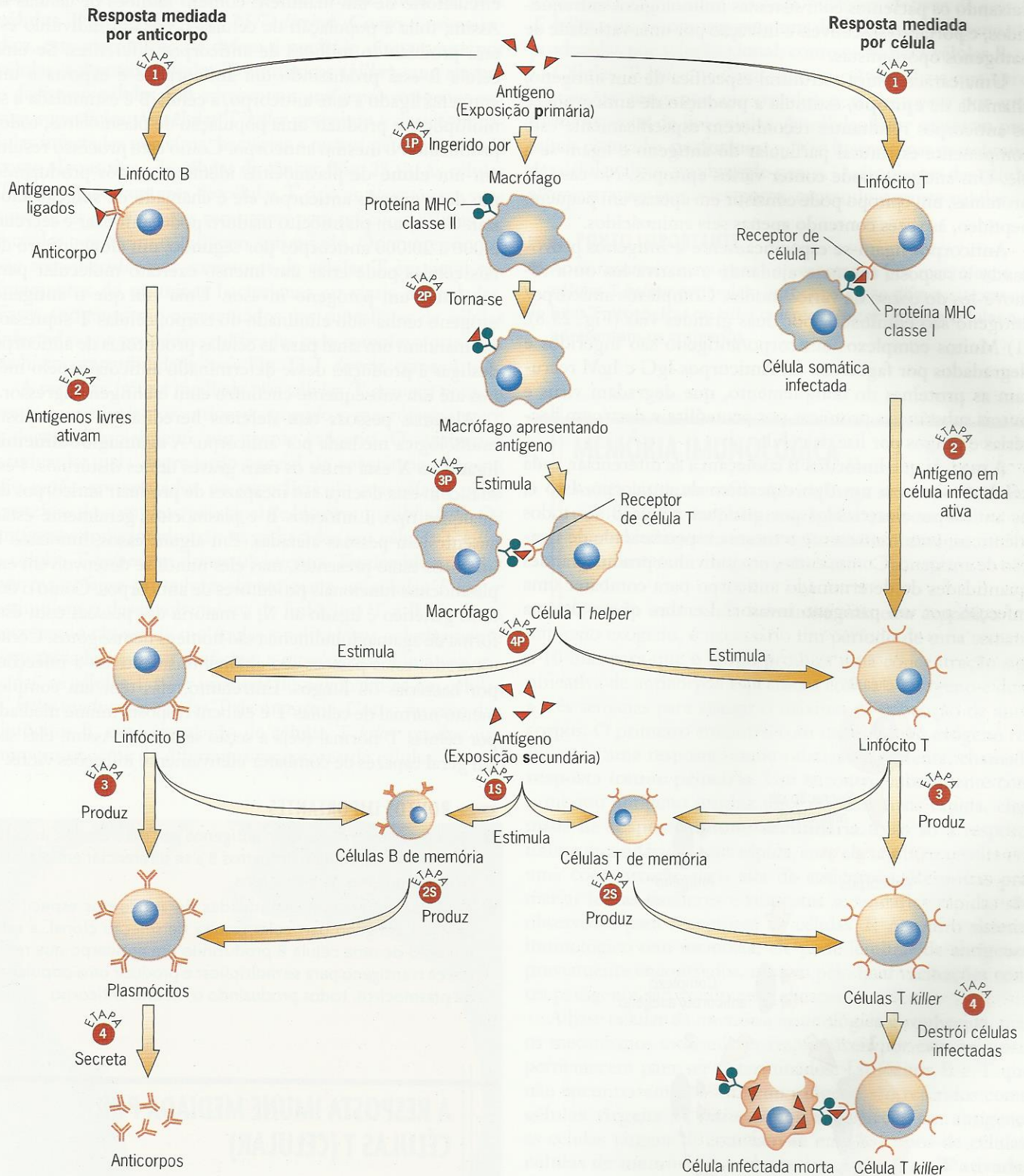
Agora que estamos familiarizados com os componentes mais importantes do sistema imunológico de mamíferos, vamos examinar as respostas imunológicas em mais detalhe.

Quando um antígeno exógeno, como uma glicoproteína na superfície de uma bactéria, entra na corrente sanguínea de um mamífero, ele dispara um mecanismo de defesa, a **resposta imunológica mediada por anticorpo**, ou **humoral**, que resulta na síntese de anticorpos. Esses anticorpos ligam-se ao antígeno com especificidade excepcional, facilitando assim a remoção da bactéria invasora do sistema circulatório. O sistema imunológico mediado por anticorpo é primariamente responsável por proteger mamíferos de infecções por bactérias e fungos e por remover vírus antes que infectem células hospedeiras. Ele envolve várias classes diferentes de células imunológicas e leva à produção de anticorpos específicos para o antígeno invasor por plasmócitos derivados de linfócitos B (Fig. 23.7, esquerda).

A resposta imune mediada por anticorpos é iniciada quando um macrófago encontra uma substância exógena tal como uma bactéria e a engloba e destrói. Antígenos de uma bactéria parcialmente degradada são ancorados na superfície do macrófago, onde são reconhecidos pelas células *T helper*. Uma vez apresentadas com antígenos pelos macrófagos, as células *T helper* ativadas estimulam linfócitos B a se diferenciar em plasmócitos, que sintetizam e secretam grandes quantidades de anticorpos, e células B de memória, que facilitam uma resposta imune mais rápida durante encontros subsequentes com o mesmo antígeno.

Em pacientes com AIDS, células *T helper* são infectadas e mortas pelo vírus da imunodeficiência humana (Cap. 18),





**Fig. 23.7** ■ Resumo da resposta imunológica em mamíferos. A resposta imunológica mediada por anticorpo (humoral) (esquerda), que é ativada por antígenos livres, é feita por linfócitos B e a resposta imunológica mediada por célula (celular) (direita), que é ativada por antígenos apresentados em células infectadas, é realizada por linfócitos T. Ambas as vias são estimuladas por células T helper ativadas produzidas em resposta à apresentação de antígenos por macrófagos (centro, etapas 1P a 4P). As respostas imunológicas rápidas que ocorrem após uma segunda ou subsequente exposição a um antígeno são mediadas por linfócitos de memória B e T que produzem rapidamente plasmócitos produtores de anticorpos e células T killer, respectivamente (centro; etapas 1S e 2S).



deixando os pacientes com sistemas imunológicos enfraquecidos, e portanto suscetíveis a infecção por uma variedade de patógenos oportunistas.

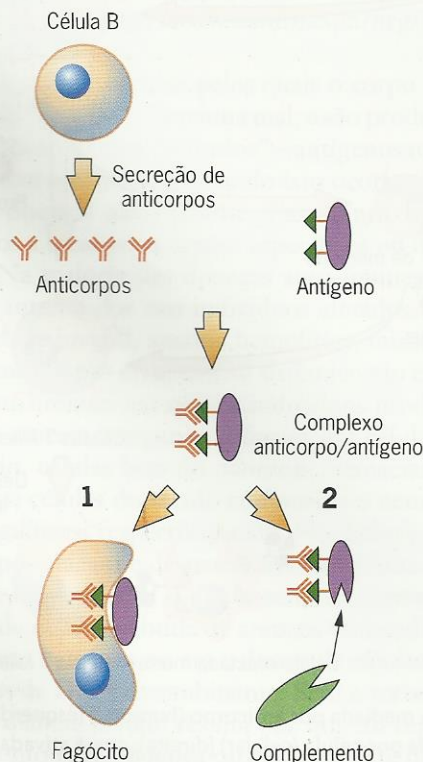
Uma característica estrutural específica de um antígeno, chamada de **epítipo**, estimula a produção de anticorpos, e os anticorpos resultantes reconhecem especificamente este componente estrutural particular do antígeno e ligam-se a ele. Um antígeno pode conter vários epítipos. No caso de proteínas, um epítipo pode consistir em apenas um pequeno peptídeo, às vezes contendo apenas seis aminoácidos.

Anticorpos ligam-se especificamente a antígenos provocando a resposta imune e ajudando a inativá-los ou a removê-los do corpo de vários modos. Complexos anticorpo/antígeno são eliminados por duas grandes vias (Fig. 23.8). (1) Muitos complexos anticorpo/antígeno são ingeridos e degradados por fagócitos. (2) Anticorpos IgG e IgM recrutam as proteínas do complemento, que degradam vírus e outras substâncias protéicas por proteólise e destroem bactérias e fungos por lise da célula.

À medida que linfócitos B começam a se diferenciar, cada célula B sintetiza um tipo específico de anticorpo. Isto é, os anticorpos sintetizados por qualquer célula B são todos idênticos, com exatamente a mesma especificidade de ligação de antígeno. Como, então, um indivíduo produz grandes quantidades de determinado anticorpo para combater uma infecção por um patógeno invasor? Lembre que o sistema

circulatório de um mamífero contém milhões de células B. Assim, toda a população de células B em um indivíduo estará produzindo milhões de anticorpos diferentes. Se uma célula B está produzindo um anticorpo e é exposta a um antígeno ligado a este anticorpo, a célula B é estimulada a se multiplicar e produzir uma população de plasmócitos, todos produzindo o mesmo anticorpo. Como este processo resulta em um **clone** de plasmócitos idênticos, todos produzindo um determinado anticorpo, ele é chamado de **seleção clonal**. Como um plasmócito maduro pode sintetizar e secretar 2.000 a 20.000 anticorpos por segundo, um clone inteiro de tais células pode criar um imenso exército molecular para combater um patógeno invasor. Uma vez que o antígeno exógeno tenha sido eliminado do corpo, células T supressoras mandam um sinal para as células produtoras de anticorpo desligar a produção desse determinado anticorpo, pelo menos até um subsequente encontro com o antígeno agressor.

Algumas pessoas têm defeitos hereditários na resposta imunológica mediada por anticorpo. A **agamaglobulinemia ligada ao X** está entre os mais graves destes distúrbios. Pessoas com esta doença são incapazes de produzir anticorpos de qualquer tipo. Linfócitos B e plasmócitos geralmente estão ausentes em pessoas afetadas. Em alguns casos, linfócitos B imaturos estão presentes, mas eles nunca se desenvolvem em plasmócitos funcionais produtores de anticorpos. Como o defeito genético é ligado ao X, a maioria das pessoas com esta forma de agamaglobulinemia são homens hemizigotos. Como esperado, estas pessoas são altamente suscetíveis a infecções por bactérias ou fungos. Entretanto, elas têm um complemento normal de células T e exibem resposta imune mediada por células T normal (veja a seção seguinte). Assim, elas são em geral capazes de combater efetivamente infecções virais.



**Fig. 23.8** ■ Complexos anticorpo/antígeno são eliminados por duas vias principais: (1) ingestão e degradação de fagócitos e (2) a atividade proteolítica das proteínas do complemento, que degradam os revestimentos protéicos de vírus e lisam patógenos celulares.

#### PONTOS IMPORTANTES

- Quando apresentadas com antígenos por macrófagos, as células T *helper* estimulam linfócitos B a se diferenciar em plasmócitos produtores de anticorpos.
- A síntese de grandes quantidades de anticorpos específicos para um antígeno particular resulta de seleção clonal, a estimulação de uma célula B produzindo um anticorpo que reconhece o antígeno para se multiplicar e produzir uma população de plasmócitos, todos produzindo o mesmo anticorpo.

### A RESPOSTA IMUNE MEDIADA POR CÉLULAS T (CELULAR)

Quando apresentadas com antígenos por macrófagos, as células T *helper* estimulam linfócitos T a se diferenciar em células T *killer* que reconhecem e lisam células que exibem esses antígenos em suas superfícies.

Além da resposta imune mediada por anticorpos, mamíferos possuem um segundo mecanismo de defesa anti-



geno-específico, chamado **resposta imune mediada por células T** ou **celular** (Fig. 23.7, *direita*). Como o nome indica, este mecanismo de defesa é mediado diretamente por células, sobretudo células T citotóxicas (*killer*), que atacam e destroem células que apresentam antígenos exógenos em suas superfícies. As células T *killer* procuram e destroem células infectadas por vírus, bactérias e outros patógenos, bem como alguns tipos de células de câncer (Fig. 23.9). As células T *killer* têm receptores de células T tipo anticorpos em suas superfícies e usam esses receptores para procurar células infectadas que levam o antígeno apropriado. Os antígenos que são reconhecidos por receptores de células T em geral são fragmentos de proteínas bacterianas ou virais degradadas. Esses antígenos são apresentados nas superfícies de células infectadas por proteínas codificadas pelo complexo principal de histocompatibilidade (veja Fig. 23.7, *direita*, etapas 1 e 2).

A resposta imune mediada por células T começa como a resposta mediada por anticorpo quando o macrófago ingere e degrada parcialmente um agente exógeno. Os antígenos produzidos pela degradação parcial da substância exógena são então apresentados na superfície do macrófago pelas proteínas do complexo principal de histocompatibilidade. Os antígenos apresentados são reconhecidos por células T *killer* e células T *helper* em desenvolvimento. Tendo sido ativadas por macrófagos portadores de antígenos, as células T *helper* estimulam o desenvolvimento de linfócitos T e B imaturos secretando classes de moléculas sinalizadoras coletivamente denominadas *citocinas*, *linfocinas* e *interleucinas*. Uma vez ativadas, as células T *killer* imaturas desenvolvem-se em células T *killer* maduras e células T de memória. Como no caso dos linfócitos B, uma população de células T *killer* produz um número enorme de diferentes receptores de células T anti-

geno-específicos. Além disso, populações efetivas de células T *killer* com receptores para um determinado antígeno são produzidas por seleção clonal, como o são para células B.

Células T *killer* maduras atacam células alvos que apresentam os antígenos apropriados e as destroem por uma forma especial de lise celular. As células *killer* secretam proteínas chamadas **perforinas**, que se inserem nas membranas das células alvos, produzindo poros. O citoplasma das células alvos extravasa por estes poros, causando a morte celular.

#### ■ PONTO IMPORTANTE

- Após reconhecer antígenos apresentados por macrófagos, as células T *helper* estimulam linfócitos T a se diferenciar em células T citotóxicas ou *killer* que reconhecem e lisam células que apresentam os antígenos.

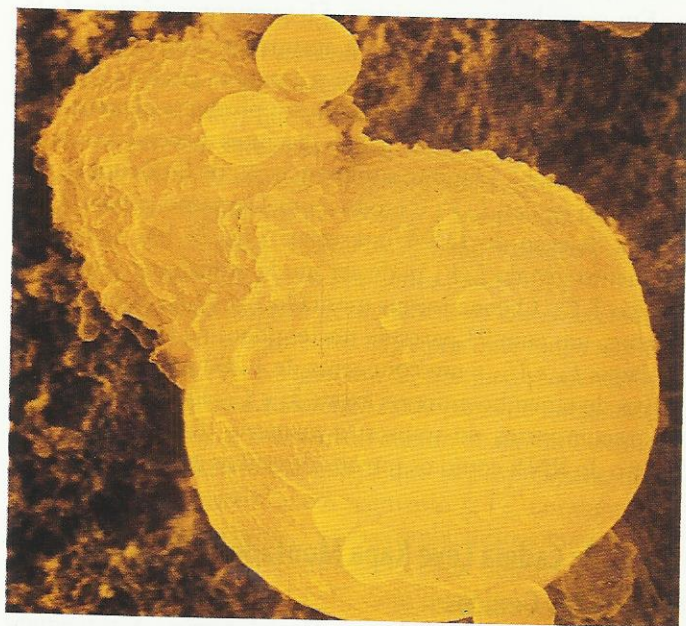
## MEMÓRIA IMUNOLÓGICA

As células de memória de vida longa são responsáveis por respostas imunológicas rápidas que ocorrem durante o segundo encontro com vírus, bactérias ou outra entidade exógena invasores.

Quando uma célula humana é primeiro exposta a um antígeno exógeno, é necessário um período de uma semana a 10 dias para que o corpo produza uma concentração significativa de anticorpos específicos contra o antígeno e duas a três semanas para atingir o máximo de produção de anticorpos. O primeiro encontro com um antígeno exógeno resulta em uma resposta imune relativamente lenta, chamada **resposta imune primária**. Um encontro subsequente com o mesmo antígeno produz uma resposta mais rápida, chamada de **resposta imune secundária**. Não só a resposta imune secundária é mais rápida, mas ela também resulta em uma concentração mais alta de anticorpos. Respostas primárias lentas similares e respostas secundárias rápidas são observadas para receptores de células T. Assim, o sistema imunológico tem memória; ele pode lembrar de antígenos previamente encontrados, motivo pelo qual vacinações contra patógenos são geralmente eficazes.

A base celular da memória imunológica é conhecida, mas os mecanismos moleculares responsáveis por essa memória permanecem para ser determinados. Linfócitos B e T que não encontraram um antígeno exógeno são referidos como **células virgens** ou *naive*. Após exposição a um antígeno, as células virgens diferenciam-se em dois tipos de células: **células de memória** e as chamadas **célula B** ou **T ativadas** (Fig. 23.7, etapa 3). As células ativadas irão multiplicar-se e diferenciar-se em plasmócitos produtores de anticorpos ou células T produtoras de receptores, respectivamente.

Em contraste com o curto tempo de vida (geralmente dias a semanas) de plasmócitos e células T *killer*, as células de memória têm vida longa, com duração de meses ou mesmo anos. Além disso, as células de memória B e T da pessoa já estão no



**Fig. 23.9** ■ Lise de uma célula alvo apresentando um antígeno exógeno por uma célula T *killer* (acima, à esquerda).



estado ativado. Elas podem, portanto, desenvolver-se em plasmócitos produtores de anticorpos e células T *killer* produtoras de receptores, respectivamente, mais rapidamente do que as células virgens. Durante uma segunda ou subsequente exposição a um antígeno, as células B e T de memória são estimuladas a se multiplicar e se diferenciar em plasmócitos e células T *killer*. Portanto, respostas imunes secundárias são mais rápidas e resultam em populações maiores de plasmócitos e células T *killer* do que respostas imunes primárias.

#### PONTO IMPORTANTE

- As células de memória de vida longa facilitam uma resposta imune mais rápida durante um segundo encontro com uma substância exógena.

## REARRANJOS GENÔMICOS DURANTE DIFERENCIAÇÃO DO LINFÓCITO B

A enorme variedade de anticorpos produzidos pelos mamíferos resulta da junção combinatória de segmentos gênicos de anticorpos por recombinação somática e por hipermutação somática durante a diferenciação de linfócitos B em plasmócitos produtores de anticorpos.

Geneticamente, o aspecto mais marcante da resposta imune é o modo pelo qual uma variedade aparentemente infinita de anticorpos podem ser sintetizados em resposta a antígenos que um animal não havia previamente encontrado. Como o genoma humano pode estocar informação genética suficiente para codificar todos esses anticorpos? A resposta a esta pergunta é simples se observarmos os detalhes moleculares. Examinamos as evidências importantes que levaram cientistas a esta resposta em Um Marco na Genética: Montagem de Genes de Anticorpos.

A informação genética que codifica cadeias de anticorpos é estocada em pedaços, e esses pedaços são juntados nas seqüências apropriadas por rearranjos genômicos que ocorrem durante o desenvolvimento das células produtoras de anticorpos do corpo. Cada cadeia de anticorpo é sintetizada usando a informação estocada como vários segmentos gênicos diferentes em cromossomos da linhagem germinativa. Os genes de anticorpos que são expressos em linfócitos B e plasmócitos são recompostos a partir desses segmentos gênicos por eventos de recombinação que ocorrem durante a diferenciação destas células a partir de células tronco na medula óssea. Durante esses rearranjos genômicos, os segmentos gênicos apropriados são unidos por processos de recombinação específicos de seqüências que deletam as regiões do DNA entre tais segmentos gênicos.

Os genes de cadeias leves lambda, genes de cadeias leves kapa e genes de cadeias pesadas são recompostos por mecanismos muito similares. A única diferença principal é que genes de cadeias leves lambda, genes de cadeias leves kapa

e genes de cadeias pesadas são montados a partir de dois, três e quatro segmentos gênicos diferentes, respectivamente. Entretanto, note que os segmentos gênicos codificantes de cadeias lambda, kapa e cadeias pesadas estão situados em três cromossomos humanos diferentes (cromossomos 22, 2 e 14, respectivamente). Assim, a montagem de genes humanos de cadeias lambda, kapa e cadeias pesadas envolve três processos de recombinação distintos.

## Genes de Cadeia Leve Lambda Montados a Partir de Dois Segmentos Gênicos

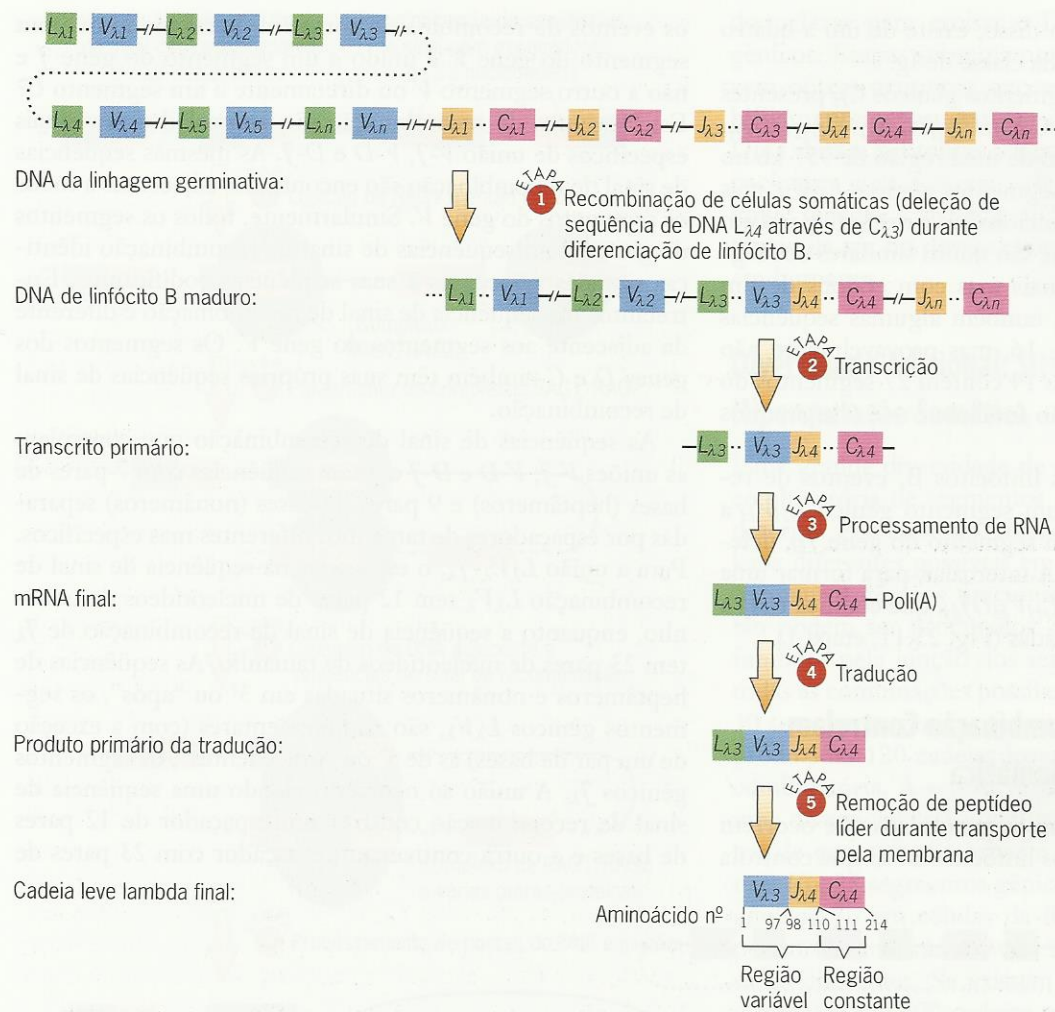
Em cromossomos da linhagem germinativa, as seqüências de DNA que codificam cadeias leves lambda estão presentes em dois tipos distintos de segmentos gênicos (Fig. 23.10). Um tipo de segmento gênico, designado  $L_\lambda V_\lambda$  (de peptídeo líder e região variável), codifica (1) o peptídeo líder hidrofóbico aminoterminal que é removido da cadeia de anticorpo após direcionar o polipeptídeo nascente através da membrana do retículo endoplasmático granular e (2) os 97 aminoácidos aminoterminal da região variável da cadeia leve lambda. O outro tipo de segmento gênico, chamado  $J_\lambda C_\lambda$  (de segmento de junção e região constante), contém a chamada **seqüência de junção** e região codificante da região constante carboxila-terminal da cadeia leve lambda. A seqüência de junção codifica os últimos 13 a 15 aminoácidos da região variável das cadeias leves lambda. Um gene completo da cadeia leve lambda é montado pela união de um segmento gênico  $L_\lambda V_\lambda$  com um segmento gênico  $J_\lambda C_\lambda$ , com a deleção do DNA intercalar, durante diferenciação do linfócito B (Fig. 23.10). O processo de montagem de gene é controlado por eventos de recombinação específicos de seqüência. As seqüências que regulam este processo são discutidas em uma seção subsequente deste capítulo.

Em seres humanos, existem 51 segmentos gênicos  $L_\lambda V_\lambda$  diferentes (21 não funcionais) agrupados perto do centrômero no cromossomo 22 e sete segmentos gênicos  $J_\lambda C_\lambda$  diferentes – quatro dos quais codificam produtos gênicos funcionais – situados um pouco mais distantes do centrômero no cromossomo 22. Durante a diferenciação de linfócitos B, um segmento gênico  $L_\lambda V_\lambda$  é unido a um segmento gênico  $J_\lambda C_\lambda$ , com a deleção do DNA intercalar (Fig. 23.10). Qualquer segmento gênico  $L_\lambda V_\lambda$  pode ser unido a qualquer segmento gênico  $J_\lambda C_\lambda$ . O gene recém-montado  $L_\lambda V_\lambda J_\lambda C_\lambda$  contém dois íntrons, um com 93 pares de nucleotídeos de tamanho entre os éxons  $L_\lambda$  e  $V_\lambda$ , e um com cerca de 1.200 pares de nucleotídeos de tamanho entre os éxons  $J_\lambda$  e  $C_\lambda$ . Estes dois íntrons são recompostos do transcrito primário durante processamento do RNA como os íntrons de outros genes.

## Genes de Cadeia Leve Kapa Montados a Partir de Três Segmentos Gênicos

A síntese de cadeias leve kapa é controlada por três segmentos gênicos diferentes: (1) um segmento gênico  $L_\kappa V_\kappa$ ,





**Fig. 23.10** ■ O controle genético das cadeias leves lambda de anticorpo humano. Cada gene de cadeia leve lambda é montado a partir de dois segmentos gênicos presentes no cromossomo 22 da linhagem germinativa durante diferenciação do linfócito B.

que codifica o peptídeo líder e o amino-terminal com 95 aminoácidos da região variável, (2) um segmento  $J_{\lambda}$ , que codifica os últimos 13 aminoácidos da região variável, e (3) um segmento gênico  $C_{\lambda}$ , que codifica a região constante carboxila-terminal. Em seres humanos, o cromossomo 2 contém 76 segmentos gênicos  $L_{\lambda}V_{\lambda}$  (embora apenas 40 sejam funcionais), cinco segmentos gênicos  $J_{\lambda}$ , e apenas um segmento gênico  $C_{\lambda}$ . Os segmentos gênicos  $J_{\lambda}$  estão situados entre os segmentos gênicos  $L_{\lambda}V_{\lambda}$  e o segmento gênico  $C_{\lambda}$ .

Em células da linhagem germinativa, os cinco segmentos  $J_{\lambda}$  são separados dos segmentos  $L_{\lambda}V_{\lambda}$  por uma longa sequência não codificante e do segmento  $C_{\lambda}$  por uma sequência não codificante de aproximadamente 2.000 pares de nucleotídeos de tamanho. Durante o desenvolvimento de um linfócito B, o gene particular de cadeia leve kapa que será expresso nesta célula é montado a partir de um segmento  $L_{\lambda}V_{\lambda}$ , um segmento  $J_{\lambda}$ , e um único segmento  $C_{\lambda}$  por recombinação somática. Durante este processo, a recombinação pode unir qualquer um dos 76 segmentos gênicos  $L_{\lambda}V_{\lambda}$  com qualquer um dos cinco segmentos  $J_{\lambda}$ , com a deleção de todo o DNA intercalar. Ela produz uma fusão  $L_{\lambda}V_{\lambda}J_{\lambda}$  que codifica toda a região variável da cadeia leve kapa. A sequência não codificante entre o

grupo de segmentos gênicos  $J_{\lambda}$  e o segmento gênico  $C_{\lambda}$  juntamente com quaisquer segmentos gênicos  $J_{\lambda}$  não deletados  $C_{\lambda}$ -proximal é retida (entre a nova fusão  $L_{\lambda}V_{\lambda}J_{\lambda}$  e o segmento  $C_{\lambda}$ ) no DNA de plasmócitos. Esta sequência inteira de DNA ( $L_{\lambda}V_{\lambda}J_{\lambda}$ -não codificante- $C_{\lambda}$ ) é transcrita, e as sequências não codificantes são removidas durante processamento de RNA, como os íntrons de qualquer outro gene eucariótico. O mRNA resultante é traduzido, e o peptídeo líder é removido do polipeptídeo kapa nascente durante seu transporte através da membrana do retículo endoplasmático granular.

### Genes de Cadeia Pesada Montados a Partir de Quatro Segmentos Gênicos

A informação genética que codifica cadeias pesadas de anticorpo é organizada em segmentos gênicos  $L_HV_H$ ,  $J_H$  e  $C_H$  análogos aos das cadeias leve kapa; mas um segmento gênico adicional, chamado *D* de diversidade, codifica 2 a 13 aminoácidos da região variável. A região variável da cadeia pesada é portanto especificada por três segmentos gênicos separados,  $L_HV_H$ ,  $J_H$  e *D*, que devem ser unidos durante o desenvol-



vimento de linfócitos B. Além disso, existe de um a quatro segmentos gênicos  $C_H$  para cada classe de Ig.

Em humanos, existem 9 segmentos gênicos  $C_H$  presentes no cromossomo 14:  $C_{H\mu}$ ,  $C_{H\delta}$ ,  $C_{H\gamma 3}$ ,  $C_{H\gamma 1}$ ,  $C_{H\alpha 1}$ ,  $C_{H\gamma 2}$ ,  $C_{H\gamma 4}$ ,  $C_{H\epsilon}$  e  $C_{H\alpha 2}$ . Além disso, dentro de uma região de 957 kb no cromossomo 14, existem 123 segmentos gênicos  $L_H V_H$ , mas apenas 44 são considerados funcionais. O restante parecem ser *pseudogenes* – seqüências que são muito similares aos segmentos gênicos  $L_H V_H$  funcionais mas com códons de fim na matriz de leitura. Existem também algumas seqüências  $L_H V_H$  nos cromossomos 15 e 16, mas provavelmente não são funcionais. O cromossomo 14 contém 27 segmentos do gene  $D$  – 25 funcionais e 2 não funcionais – e 6 segmentos gênicos  $J_H$  (Fig. 23.11).

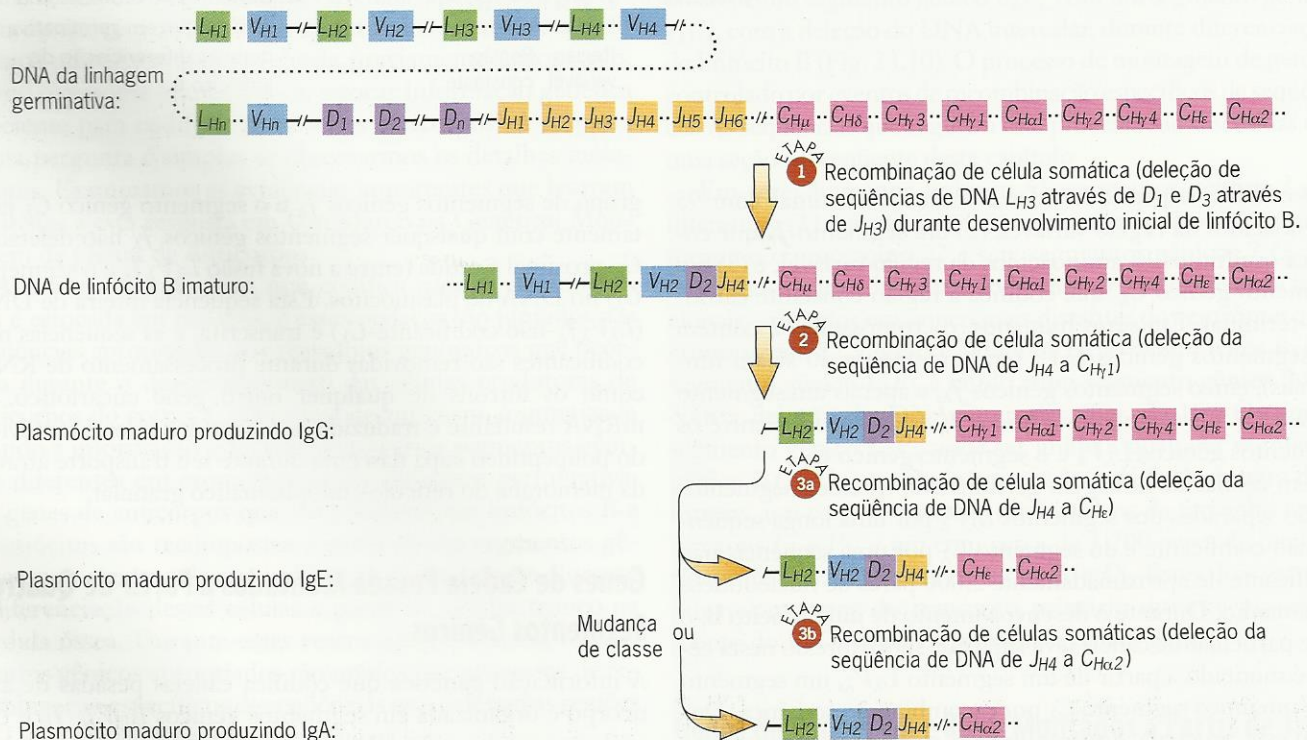
Durante diferenciação dos linfócitos B, eventos de recombinação somática unem um segmento gênico  $L_H V_H$  a um segmento do gene  $D$  e um segmento do gene  $J_H$ , deletando duas seqüências de DNA intercalar, para formar uma seqüência contínua de DNA,  $L_H V_H D J_H$ , que codifica toda a região variável das cadeias pesadas (Fig. 23.11, etapa 1).

### Seqüências de Sinal de Recombinação Controlam Eventos de Recombinação Somática

Como são regulados os rearranjos genômicos que ocorrem durante o desenvolvimento dos linfócitos B? O que controla

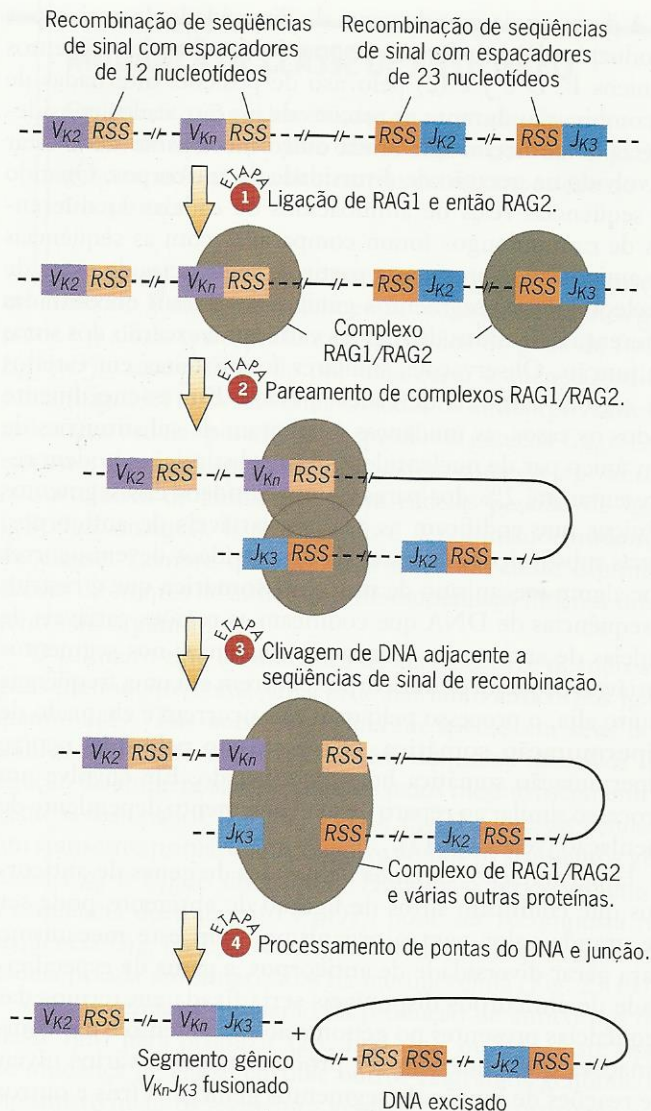
os eventos de recombinação somática de tal modo que um segmento do gene  $V$  é unido a um segmento do gene  $J$  e não a outro segmento  $V$  ou diretamente a um segmento  $C$ ? Estes eventos de recombinação são controlados por sinais específicos de união  $V-J$ ,  $V-D$  e  $D-J$ . As mesmas seqüências de sinal de recombinação são encontradas adjacentes a todos os segmentos do gene  $V$ . Similarmente, todos os segmentos do gene  $J$  têm seqüências de sinal de recombinação idênticas situadas adjacentes a suas seqüências codificantes. Entretanto, sua seqüência de sinal de recombinação é diferente da adjacente aos segmentos do gene  $V$ . Os segmentos dos genes  $D$  e  $C$  também têm suas próprias seqüências de sinal de recombinação.

As seqüências de sinal de recombinação que controlam as uniões  $V-J$ ,  $V-D$  e  $D-J$  contêm seqüências com 7 pares de bases (heptâmeros) e 9 pares de bases (nonâmeros) separadas por espaçadores de tamanhos diferentes mas específicos. Para a união  $L_\lambda V_\lambda-J_\lambda$ , o espaçador na seqüência de sinal de recombinação  $L_\lambda V_\lambda$  tem 12 pares de nucleotídeos de tamanho, enquanto a seqüência de sinal de recombinação de  $J_\lambda$  tem 23 pares de nucleotídeos de tamanho. As seqüências de heptâmeros e nonâmeros situadas em 3' ou "após", os segmentos gênicos  $L_\lambda V_\lambda$ , são complementares (com a exceção de um par de bases) às de 5' ou "precedentes", os segmentos gênicos  $J_\lambda$ . A união só ocorrerá quando uma seqüência de sinal de recombinação contiver um espaçador de 12 pares de bases e a outra contiver um espaçador com 23 pares de



**Fig. 23.11** ■ O controle genético das cadeias pesadas de anticorpo humano. Cada gene de cadeia pesada é montado a partir de quatro segmentos gênicos diferentes situados no cromossomo 14 por eventos de recombinação somática que ocorrem durante diferenciação do linfócito B.  $C_{H\mu}$ ,  $C_{H\delta}$  e  $C_{H\epsilon}$  codificam as regiões constantes de cadeia pesada de IgM, IgD e IgE, respectivamente.  $C_{H\gamma 3}$ ,  $C_{H\gamma 1}$ ,  $C_{H\gamma 2}$  e  $C_{H\gamma 4}$  codificam regiões constantes de cadeia pesada de quatro formas bem correlatas de IgG,  $C_{H\alpha 1}$  e  $C_{H\alpha 2}$  especificam cadeias pesadas de duas formas de IgA.





**Fig. 23.12** ■ Modelo abreviado de junção  $V_{\lambda}-J_{\lambda}$ . O processo de junção é mediado pela ligação específica de RAG1 e RAG2 às sequências de sinal de recombinação (RSS) flanqueadoras dos segmentos gênicos  $V_{\lambda}$  e  $J_{\lambda}$ . As RSS adjacentes a segmentos  $V_{\lambda}$  contêm espaçadores com 12 nucleotídeos; aquelas adjacentes a segmentos  $J_{\lambda}$  contêm espaçadores com 23 nucleotídeos. O complexo proteico RAG1/RAG2 catalisa recombinação apenas quando uma RSS contém um espaçador de 12 nucleotídeos e outra contém um espaçador de 23 nucleotídeos. Detalhes do processo de junção, incluindo outras proteínas necessárias, foram omitidos.

bases. Esta exigência é criada pelas proteínas RAG1 (gene I ativador de recombinação) e RAG2 que medeiam os processos de recomposição do DNA (Fig. 23.12). Sequências de sinal de recombinação muito similares controlam a união de  $L_HV_H-D$  e  $D-J_H$ .

A união mediada por RAG1/RAG2 dos segmentos gênicos  $V_{\lambda}L_{\lambda}$  e  $J_{\lambda}$  começa com a ligação de RAG1 às duas sequências de sinal de recombinação. RAG2 então se junta ao complexo, e as duas proteínas funcionam como uma en-

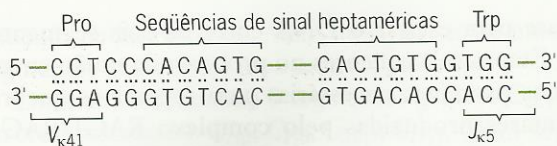
donuclease para excisar o DNA entre os dois segmentos gênicos. Etapas subsequentes no processo de união envolvem outras proteínas necessárias para reparar as quebras bifilamentares produzidas pelo complexo RAG1/RAG2. Uma dessas proteínas, chamada Artemis, é defeituosa em pacientes com a doença de imunodeficiência combinada grave sensível a radiação (RS-SCID), destacando a importância da via no desenvolvimento do sistema imunológico em humanos.

### Diversidade Adicional: Sítios de União Variável e Hipermutação Somática

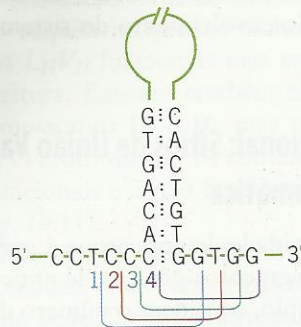
Uma grande diversidade de anticorpos é gerada pela união combinatória de segmentos gênicos de anticorpos como já descrito. Por exemplo, considere o número de cadeias leves kapa diferentes possíveis em seres humanos. Se existem 40 segmentos gênicos funcionais  $L_{\lambda}V_{\lambda}$  e 5 segmentos  $J_{\lambda}$ , então podem ser produzidos 200 segmentos gênicos  $L_{\lambda}V_{\lambda}J_{\lambda}$  fundidos pela junção dos segmentos gênicos  $L_{\lambda}V_{\lambda}$  e  $J_{\lambda}$  em todas as combinações possíveis. Similarmente, considerando 30 segmentos gênicos funcionais  $L_{\lambda}V_{\lambda}$  e 4  $J_{\lambda}C_{\lambda}$ , podem ser produzidas 120 cadeias leves lambda diferentes por junção combinatória. A região variável de cadeias pesadas fornece uma diversidade adicional por causa dos múltiplos segmentos de genes  $D$ . Se existem 44 segmentos gênicos funcionais  $V_H$ , 25 segmentos gênicos  $D$  funcionais, e 6 segmentos gênicos  $J_H$  em células da linhagem germinativa humana, podem ser montadas 6.600 regiões variáveis de cadeias pesadas diferentes. Se existem 320 (200 + 120) cadeias leves diferentes e 6.600 cadeias pesadas diferentes, então  $320 \times 6.600$ , ou 2.112.000, anticorpos diferentes podem ser produzidos por junção combinatória apenas. Claramente, a junção combinatória de segmentos gênicos de anticorpos produz uma vasta quantidade de diversidade de ligação de antígeno. Entretanto, análises moleculares de anticorpos e genes de anticorpos mostraram que a diversidade adicional é gerada por outros mecanismos.

Uma comparação da diversidade de sequências de aminoácidos presente em moléculas de anticorpo com a prevista pelas sequências de segmentos gênicos que codificam tais anticorpos revela que há mais variação em sequências de aminoácidos nas junções  $V-J$  do que é previsto pelas sequências de nucleotídeos. Foi demonstrado que parte desta diversidade adicional resulta do uso de sítios alternativos de recombinação durante os eventos de junção que estão envolvidos na montagem final de genes de anticorpo maduro. Um exemplo do uso de sítios alternativos de junção de segmentos gênicos  $V_{\lambda}$  e  $J_{\lambda}$  no camundongo é ilustrado na Fig. 23.13. Durante a união dos segmentos  $V_{\lambda 41}$  e  $J_{\lambda 5}$ , foi mostrado que eventos de recombinação ocorrem entre quatro posições de nucleotídeos adjacentes nos sítios de junção. Como ilustrado na Fig. 23.13d, estes eventos de recombinação produzem quatro sequências diferentes de nucleotídeos que codificam três aminoácidos distintos na

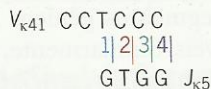




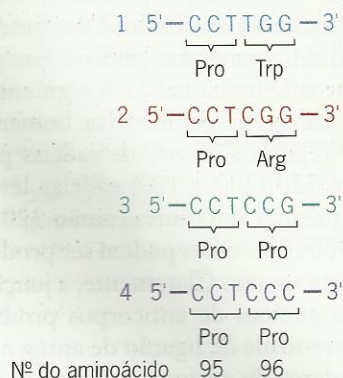
(a) Sequências de nucleotídeos de segmentos gênicos  $V_{K41}$  e  $J_{K5}$  no camundongo.



(b) Papel hipotético das sequências de sinal na justaposição dos segmentos gênicos.



(c) Sítios alternativos de recombinação somática.



(d) Sequências de nucleotídeos resultantes e os dipeptídeos que elas codificam.

**Fig. 23.13** ■ A diversidade de anticorpos na junção  $V_{\lambda}-J_{\lambda}$  é produzida por variação na posição exata da reação de junção.

posição 96 na cadeia kapa leve de camundongo. Como o aminoácido 96 ocorre em uma região da cadeia de anticorpo que está envolvida na ligação de antígeno, eventos de junção alternados  $V-J$  deste tipo contribuem significativamente para a grande diversidade de especificidade de anticorpos que é observada em mamíferos. Eventos de junção alternados semelhantes foram documentados para reações de junção  $V_{\lambda}-J_{\lambda}$  e  $V_H-D-J_H$ . Além disso, as reações de junção são imprecisas, em geral deletando ou inserindo de um a alguns nucleotídeos.

A despeito da grande gama de diversidade de anticorpos produzidos (1) pela junção de grandes famílias de segmentos gênicos  $V$ ,  $D$  e  $J$  e (2) pelo uso de posições alternadas de recombinação durante as reações de junção, dados consideráveis demonstram que ainda outro mecanismo deve estar envolvido na geração de diversidade de anticorpos. Quando as sequências reais de aminoácidos de cadeias  $\lambda_1$  diferentes de camundongos foram comparadas com as sequências de aminoácidos previstas a partir de sequências de pares de nucleotídeos de segmentos gênicos  $\lambda_1$ , foram encontradas diferenças nos sítios de regiões variáveis à exceção dos sítios de junção. Observações similares foram feitas em estudos de regiões variáveis de cadeia pesada. Em essencialmente todos os casos, as mudanças resultaram de substituições de um único par de nucleotídeos. Tais substituições podem representar até 2% dos pares de nucleotídeos dos segmentos gênicos que codificam as regiões variáveis de anticorpos. Essas substituições de pares de nucleotídeos devem ocorrer por algum mecanismo de **mutação somática** que é restrito a sequências de DNA que codificam as regiões variáveis de cadeias de anticorpos. Como tais mudanças nos segmentos variáveis de genes de anticorpos ocorrem em uma frequência muito alta, o processo pelo qual elas ocorrem é chamado de **hipermutação somática**. O mecanismo pelo qual ocorre hipermutação somática hoje é conhecido. Ele envolve um processo similar ao reparo de malpareamento dependente de metilação (veja Cap. 13).

Hipermutação somática de regiões de genes de anticorpos que codificam sítios de ligação de antígenos pode ser de grande valor para o organismo. Sem este mecanismo para gerar diversidade de anticorpos, a gama de especificidade de anticorpos disponíveis seria fixada em termos das sequências presentes no genoma ao nascimento e das combinações que poderiam ser produzidas pelos vários níveis de reações de junção de segmentos gênicos. Vírus e outros patógenos estão constantemente evoluindo e produzindo novas variantes com novos determinantes antigênicos. Para proporcionar uma defesa adequada contra a mudança de composição antigênica desses patógenos, o sistema imunológico também deve ser capaz de responder rapidamente a tais mudanças. Que modo melhor de fornecer esta salvaguarda do que capacitar genes de anticorpos com seu próprio mecanismo – hipermutação somática – para rápida adaptação a novos antígenos que podem evoluir no futuro?

### PONTOS IMPORTANTES

- O grande número de diferentes anticorpos produzidos por mamíferos resulta amplamente da junção combinatória de segmentos gênicos de anticorpos por recombinação somática durante a diferenciação de linfócitos B em plasmócitos produtores de anticorpos.
- A diversidade adicional de anticorpos é produzida por variabilidade nos sítios de junção de segmentos gênicos e por alta frequência de mutação – hipermutação somática – nas sequências de DNA que codificam sítios de ligação de antígenos.



## MUDANÇA DE CLASSE DE ANTICORPO

Os linfócitos B podem mudar da produção de uma classe de anticorpo para outras classes por rearranjo do genoma ou por vias alternativas de recomposição do transcrito.

Na época em que a síntese de anticorpos começa em um linfócito B em desenvolvimento, todos os segmentos gênicos  $C_H$  ainda estão presentes, separados do gene de fusão  $L_HV_HD\bar{J}_H$  recém-formado por uma curta sequência codificante (Fig. 23.11). Neste estágio, todos os anticorpos sintetizados têm cadeias pesadas IgM (produtos gênicos  $C_{H\mu}$ ). Entretanto, alguns plasmócitos logo podem começar a produzir outras classes de anticorpos (com cadeias pesadas de IgA, IgD, IgE ou IgG). Este processo é denominado **mudança de classe**. Lembre que, como já discutido neste capítulo, classes diferentes de anticorpos desempenham funções diferentes no organismo.

O primeiro tipo de mudança de classe ocorre quando alguns plasmócitos começam a produzir tanto anticorpos IgM quanto IgD. Tais anticorpos diferem apenas em seus domínios de função efetora; eles têm domínios idênticos de ligação de antígenos, especificados pelos mesmos segmentos gênicos fusionados  $V_L\bar{J}_L$  ou  $V_L\bar{J}_L$  e  $V_HD\bar{J}_H$ . Nestas células, um transcrito primário que envolve ambos os segmentos gênicos  $C_{H\mu}$  e  $C_{H\delta}$  é sintetizado. Durante o processamento, a sequência do transcrito  $V_HD\bar{J}_H$  pode ser recomposta na sequência  $C_{H\mu}$  ou  $C_{H\delta}$ . Como resultado, ambos os tipos de cadeia pesada são sintetizadas na mesma célula (Fig. 23.14).

Outras células B mudarão da produção de anticorpos da classe IgM para produzir IgA, IgD, IgE ou IgG. Este tipo de mudança de classe envolve outros rearranjos genômicos durante os quais os segmentos gênicos  $C_H$  mais próximos aos segmentos gênicos  $L_HV_HD\bar{J}_H$  previamente unidos são deletados (Fig. 23.11, etapa 3). A classe de anticorpos produzida

após mudança de classe depende de que região codificante  $C_H$  permanece justaposta à sequência codificante  $L_HV_HD\bar{J}_H$  previamente fusionada.

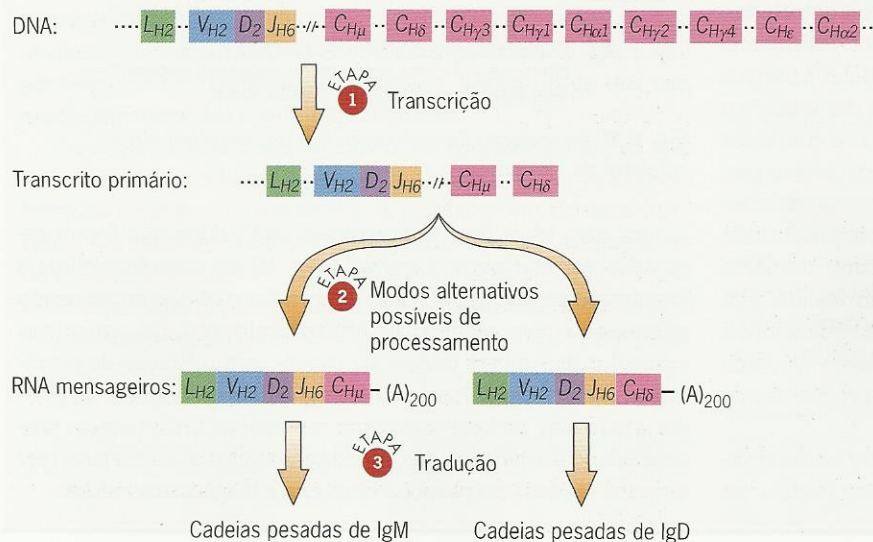
### PONTO IMPORTANTE

- Linfócitos B em desenvolvimento podem mudar da produção de anticorpos IgM para outras classes de anticorpos por rearranjos genômicos ou por vias alternativas de recomposição de transcrito.

## MONTAGEM DE GENES DE RECEPTORES DE CÉLULAS T POR RECOMBINAÇÃO SOMÁTICA

Genes receptores de células T são montados a partir de segmentos gênicos durante diferenciação do linfócito T por rearranjos genômicos similares aos envolvidos na produção de anticorpos.

Como as cadeias pesadas de anticorpos, as duas cadeias polipeptídicas de receptores de células T são codificadas por segmentos gênicos  $L-V$ ,  $D$ ,  $\bar{J}$  e  $C$ . As regiões variáveis de proteínas de receptores de células T são codificadas por vários segmentos  $L-V$ ,  $D$  e  $\bar{J}$ ; as regiões constantes são codificadas por um pequeno número de segmentos do gene  $C$ . Os genes receptores de células T são montados por rearranjos genômicos que ocorrem durante a diferenciação de linfócitos T a partir de células tronco, como no caso de genes de anticorpos em linfócitos B em desenvolvimento. As proteínas receptoras  $\alpha$  e  $\beta$  são codificadas por segmentos gênicos que são dispostos em grupos como os segmentos gênicos que codificam cadeias de anticorpos. Em humanos, os grupos de segmentos gênicos  $\alpha$  e  $\beta$  estão situados nos cromossomos 14 e 7, respectivamente.



**Fig. 23.14** ■ Mudança de classe de anticorpo pode ocorrer por vias alternadas de recomposição do transcrito. As duas vias de recomposição mostradas resultam na produção de cadeias pesadas de IgM e IgD pelo mesmo plasmócito.



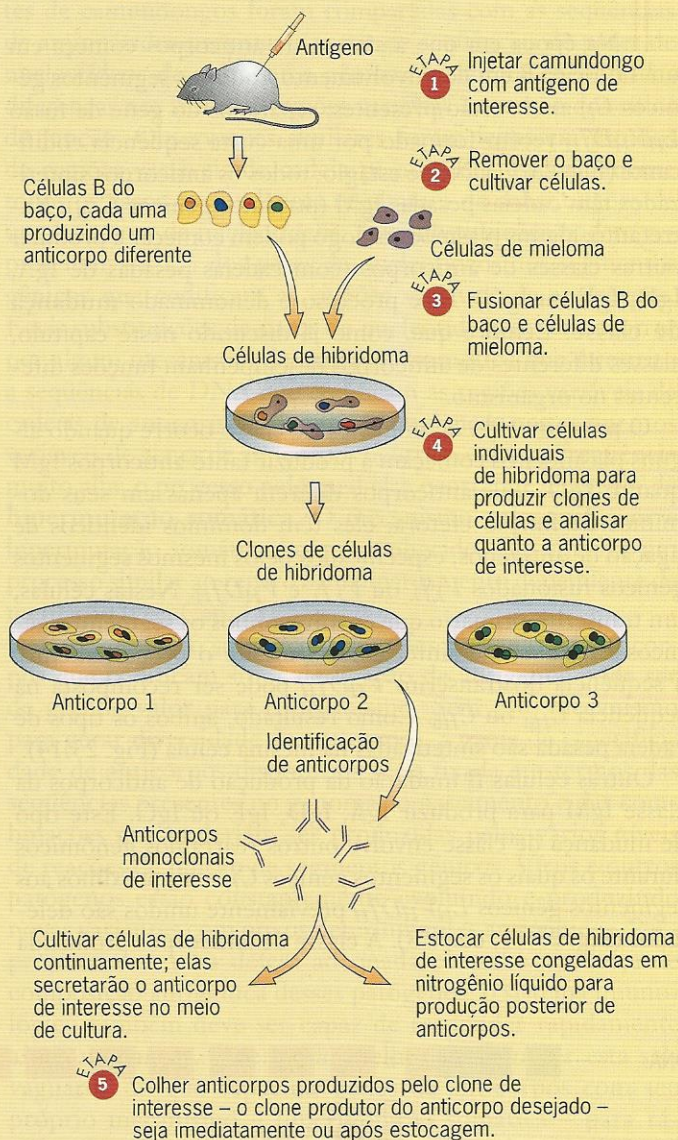


## ENFOQUE TÉCNICO: Anticorpos Monoclonais: Sondas Poderosas para Detectar Proteínas e Outros Antígenos

Além de defender o corpo contra vírus, bactérias, fungos e outras substâncias exógenas, os anticorpos servem a cientistas como poderosas ferramentas de pesquisa que podem ser usadas para identificar proteínas específicas ou outras moléculas antigênicas. Por muitos anos, o problema do uso de anticorpos como sondas moleculares foi a tarefa difícil de purificar aqueles específicos para o antígeno de interesse. Em 1975, Cesar Milstein e Georges Köhler desenvolveram um procedimento que resolveu este problema. Eles descobriram como produzir clones de culturas de células imortais secretoras de anticorpos. Eles fusionaram linfócitos B com células cancerosas chamadas mielomas para produzir células somáticas híbridas. As células híbridas resultantes, denominadas **hibridomas**, eram imortais. Além disso, elas podiam ser clonadas e triadas quanto à produção dos anticorpos de interesse. Uma vez identificadas, as linhagens celulares desejadas de hibridoma podiam ser usadas para produzir essencialmente quantidades ilimitadas de **anticorpos monoclonais** puros, todos anticorpos reconhecendo o mesmo epítipo antigênico.

O procedimento usado para produzir anticorpos monoclonais é mostrado na Fig. 1. A primeira etapa é injetar um camundongo com o antígeno de interesse. Isto irá estimular o sistema imunológico do camundongo a começar a produzir anticorpos específicos contra o antígeno. A segunda etapa é remover o baço do camundongo injetado e colocar as células do baço, muitas das quais são linfócitos B, em um meio de cultura. A terceira etapa envolve fusão de células B com células de mieloma de crescimento rápido para produzir células de hibridoma. Esta é a etapa principal na produção de anticorpos monoclonais. Cada célula B produzirá um único tipo de anticorpo devido a exclusão alélica. Entretanto, não podemos apenas clonar células B porque elas viverão apenas alguns dias, seja *in vitro* ou *in vivo*. Células cancerosas, tais como os mielomas, são imortais em cultura. Elas se dividem rápida e continuamente, exibindo crescimento e divisão descontrolados. Mas as células de mieloma não produzem anticorpos. Felizmente, as células híbridas adquirem a imortalidade das células parentais de mieloma e a capacidade de produzir anticorpos dos linfócitos B parentais. Células híbridas parentais são separadas e cultivadas para produzir clones de hibridomas produtores de anticorpos. Cada clone é analisado quanto à produção do anticorpo de interesse. Uma vez identificados, os hibridomas que estão produzindo os anticorpos desejados podem ser cultivados continuamente ou estocados depois de congelados em nitrogênio líquido para uso posterior. Assim, o procedimento de anticorpos monoclonais permite-nos produzir uma quantidade essencialmente ilimitada de um anticorpo importante e estocar as células produtoras de anticorpo para uso muitos anos depois.

Considerando a sua especificidade fenomenal, anticorpos monoclonais são valiosas ferramentas de pesquisa. Eles podem ser



**Fig. 1** ■ A produção de anticorpos monoclonais por clones celulares de hibridoma.

usados para identificar polipeptídeos individuais que foram separados por eletroforese em gel (Cap. 15) ou para determinar a localização de macromoléculas em tecidos e células combinando microscopia com técnicas de imunolocalização. Os anticorpos monoclonais também podem ser usados na purificação de proteínas por imunoprecipitação e no diagnóstico de doenças adquiridas e herdadas. Embora anticorpos monoclonais não tenham proporcionado a esperança de cura do câncer, eles demonstraram ser recursos valiosos em pesquisa biológica e diagnóstico médico.





## UM MARCO NA GENÉTICA: Montagem de Genes de Anticorpos

Nos anos 1950 e 1960, já sabíamos que animais vertebrados podiam produzir um número enorme (provavelmente milhões) de anticorpos diferentes. Se cada um desses anticorpos fosse o produto de um gene diferente, seria necessário uma grande proporção do genoma para codificar o vasto repertório de anticorpos. Entretanto, nós também sabíamos que o genoma tinha que dirigir a síntese de uma grande gama de enzimas, proteínas estruturais e moléculas de RNA. Os genomas de vertebrados não parecem grandes o suficiente para realizar ambas as funções. Uma possibilidade era que cada classe de cadeia de anticorpos – cadeias leves lambda, cadeias leves kappa e cadeias pesadas – é codificada por um único gene estrutural e que a diversidade é gerada por alta frequência de mutações somáticas nesse gene. Entretanto, se fosse o caso, por que uma alta taxa de mutação somática ocorre nestes genes e não em outros?

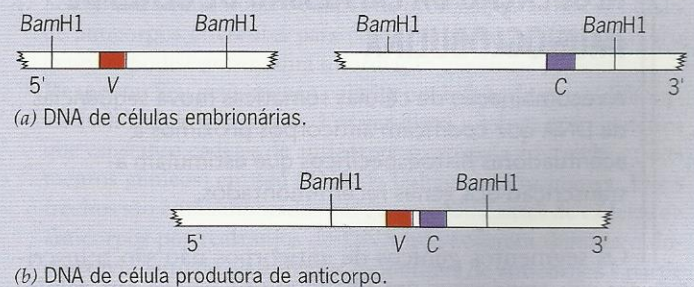
No começo dos anos 1960, foi demonstrado que anticorpos contêm quatro polipeptídeos, duas cadeias leves idênticas e duas cadeias pesadas idênticas (veja Fig. 23.3). Assim, se tivéssemos alguns genes codificando cadeias leves e outros genes codificando cadeias pesadas, seus produtos poderiam ser combinados de vários modos para produzir muitos anticorpos diferentes. Logo depois, foi demonstrado que cada anticorpo continha uma região variável em seu terminal amino e uma região constante em seu terminal carboxila. A estrutura singular das cadeias polipeptídicas em anticorpos levou William Dreyer e Claude Bennett a propor que as regiões variáveis e constantes de uma cadeia de anticorpo são codificadas por genes diferentes.<sup>1</sup> Sua proposta foi referida como a “hipótese de dois genes–um polipeptídeo”. Eles também propuseram que as regiões constantes de cadeias de anticorpo são codificadas por um único gene (o gene C) e as regiões variáveis são especificadas por milhares de genes distintos (os genes V). Finalmente, eles sugeriram que a recombinação genética moveu um dos genes V para perto de um gene C durante o desenvolvimento de plasmócitos produtores de anticorpos. Suas hipóteses receberam pouco apoio na época, em parte porque os genomas eram considerados como estruturas muito estáveis e a recombinação parecia ocorrer durante a meiose quando os cromossomos homólogos eram pareados. Os rearranjos genômicos foram considerados anormais. Acreditava-se que eles produziam mutações e fenótipos anormais.

Os primeiros resultados definitivos de apoio à hipótese de dois genes–um polipeptídeo foram publicados em 1976 por Nobumichi Hozumi e Susumu Tonegawa (Fig. 1), pesquisadores do Basel Institute for Immunology na Suíça. Eles demonstraram que as seqüências de DNA codificantes das regiões variáveis e constantes das cadeias leves kappa (uma das duas classes de cadeias leves de anticorpos) estavam presentes no mesmo fragmento de *Bam*HI do DNA genômico em células produtoras de anticorpos, mas estavam em frag-

mentos diferentes de *Bam*HI do DNA em células embrionárias (Fig. 2).<sup>2</sup> Tonegawa e colaboradores subsequentemente mostraram que isto também era verdade para seqüências de DNA codificantes das regiões variáveis e constantes da outra classe (lambda) de cadeias



**Fig. 1** ■ Susumu Tonegawa, o primeiro a demonstrar que genes codificantes de cadeias de anticorpo são recompostos juntos a partir de segmentos gênicos por rearranjos genômicos durante o desenvolvimento de plasmócitos produtores de anticorpos. Tonegawa recebeu o Prêmio Nobel de Fisiologia ou Medicina de 1987 em reconhecimento à importância de seu trabalho.



**Fig. 2** ■ A demonstração de Hozumi e Tonegawa de que seqüências de DNA codificantes das regiões variáveis e constantes de cadeias leves kappa estão em diferentes fragmentos de restrição de *Bam*HI no DNA genômico de células embrionárias (a), mas estão próximas no mesmo fragmento de *Bam*HI no DNA genômico de células produtoras de anticorpos (b).

<sup>1</sup>Dreyer, W. J. e J. C. Bennett. 1965. The molecular basis of antibody formation: A paradox. *Proc. Natl. Acad. Sci. U.S.A.* 54: 864-869.

<sup>2</sup>Hozumi, N. e S. Tonegawa. 1976. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc. Natl. Acad. Sci. U.S.A.* 73: 3628-3632.





## UM MARCO NA GENÉTICA (continuação)

leves<sup>3</sup> e cadeias pesadas.<sup>4</sup> Tonegawa recebeu o Prêmio Nobel de Fisiologia ou Medicina de 1987 por seu trabalho. Philip Leder e Leroy Hood e colaboradores também forneceram evidências importantes para a geração da diversidade de anticorpos pela junção combinatória de segmentos gênicos de anticorpos.

Hoje sabemos que genes de cadeias leves kapa e lambda são recompostos juntos de dois e três segmentos gênicos diferentes, respectivamente (veja Figs. 23.10 e 14.26) e que genes de cadeias pesadas são montados a partir de quatro segmentos gênicos diferentes (veja Fig. 23.11). Além disso, os genes codificantes de proteínas receptoras nas superfícies de linfócitos T são produzidos por rearranjos genômicos similares em células precursoras. Assim, embora a hipótese de dois genes-um polipeptídeo de Dreyer e

Bennett parecesse muito radical quando foi proposta, hoje sabemos que foi uma proposta bem engenhosa.

### QUESTÕES PARA DISCUSSÕES

1. Os resultados de Tonegawa e outros provaram que as principais características da "hipótese dois genes-um polipeptídeo" de Dreyer e Bennett estão corretas. Mas esta hipótese teve pouco apoio da comunidade científica quando foi proposta em 1965. Isto é típico de idéias novas em ciência? Os cientistas são muito conservadores? Como novos conceitos se tornam estabelecidos na ciência?
2. Em virtude da sua capacidade de produzir anticorpos contra milhões de antígenos diferentes, nosso sistema imunológico protege-nos de um grande número de vírus, bactérias e outros microrganismos patogênicos. Por que nosso sistema imunológico não nos protege do vírus da imunodeficiência humana (HIV) que causa a síndrome da imunodeficiência adquirida (AIDS)?

<sup>3</sup>Tonegawa, S., C. Brack, N. Hozumi e V. Pirrotta. 1977. Organization of immunoglobulin genes. *Cold Spring Harbor Symp. Quant. Biol.* 42: 921-931.

<sup>4</sup>Sakano, H., R. Maki, Y. Kurosawa, W. Roeder e S. Tonegawa. 1979. Domains and the hinge region of an immunoglobulin heavy chain are encoded in separate DNA segments. *Nature* 277:627-633.

### PONTO IMPORTANTE

- Os genes que codificam as proteínas receptoras de células T killer são montados a partir de segmentos gênicos durante diferenciação do linfócito T por rearranjos genômicos análogos aos envolvidos na montagem de genes de anticorpos.

## REGULAÇÃO DA EXPRESSÃO DE GENES DE IMUNOGLOBULINA

A recombinação de células somáticas move seqüências de DNA que codificam anticorpos próximos a acentuadores histoespecíficos que estimulam a transcrição dos genes recém-montados.

Os segmentos gênicos de anticorpos não são transcritos, ou são transcritos em níveis muito baixos em células da linhagem germinativa. Em plasmócitos, 10 a 20% das moléculas de mRNA são transcritos de genes de anticorpos. O que ativa a transcrição de genes de anticorpos durante ou imediatamente após sua montagem? Por que só um tipo de anticorpo é produzido por um determinado plasmócito?

### Exclusão Alélica: Apenas um Rearranjo Funcional Por Célula

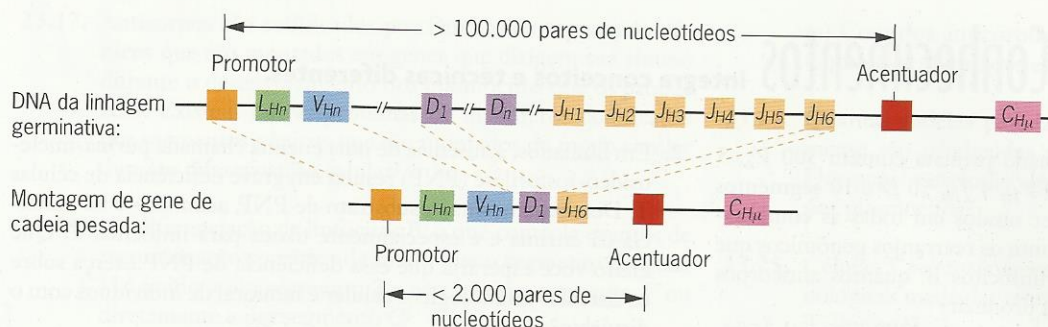
Cada plasmócito produz apenas um tipo de anticorpo. Por quê? As células de mamíferos são diplóides; elas levam dois

conjuntos de informação genética que codificam cada uma das cadeias de anticorpos. Entretanto, apenas um rearranjo genômico produtivo de cadeia leve que codifica seqüências e um rearranjo genômico produtivo de cadeia pesada que codifica seqüências ocorre durante diferenciação de cada linfócito B. Este fenômeno é chamado **exclusão alélica** porque um dos "alelos" é excluído de ser expresso. Como? Por quê? No momento, geneticistas não sabem. Eles sabem que ambos os "alelos" são, às vezes, rearranjados na mesma célula, mas, nesses casos, apenas um "alelo" sofreu um rearranjo produtivo em um gene funcional. O outro "alelo" é sempre apenas parcialmente montado ou contém deleções de seqüências codificantes essenciais. Uma consequência prática valiosa da exclusão alélica é que ela permite a pesquisadores obter grandes quantidades de um anticorpo que se liga a um único epítipo, o chamado **anticorpo monoclonal** (veja Enfoque Técnico: Anticorpos Monoclonais: Sondas Poderosas para Detectar Proteínas e Outros Antígenos), sem usar as trabalhosas metodologias laboratoriais de purificação de proteínas.

### Transcrição de Genes de Cadeia Pesada: Um Acentuador Histoespecífico

Por que mecanismo são ativados os genes de anticorpo durante sua montagem a partir de segmentos gênicos? No caso de genes de cadeia pesada, a resposta é que o processo de montagem leva os promotores situados antecedentes aos segmentos gênicos  $L_HV_H$  para a área de influência de um acentuador forte (Cap. 21) situado no íntron entre os seg-





**Fig. 23.15** ■ Ativação da transcrição de genes de cadeia pesada por movimento de um promotor de gene de cadeia pesada (boxe laranja) para a proximidade de um acentuador histoespecífico (boxe vermelho) durante montagem de genes em linfócitos B.

mentos gênicos  $J_H$  e o segmento gênico  $C_{H\mu}$  (Fig. 23.15). Cada segmento gênico  $L_HV_H$  contém um promotor antecedente. Antes dos eventos de rearranjo genômicos que resultam na síntese de cadeia pesada, o promotor mais próximo de  $L_HV_H$  dista mais de 100.000 pares de nucleotídeos do acentuador (Fig. 23.15, em cima), e os acentuadores têm pouco efeito sobre a transcrição iniciada nos promotores que estão distantes. Durante diferenciação de células B, o processo de montagem do gene de cadeia pesada (Fig. 23.11) leva o promotor do segmento gênico  $L_HV_H$  mais próximo e o acentuador para dentro de 2.000 pares de nucleotídeos um do outro (Fig. 23.15, embaixo). O acentuador agora ativa a transcrição do promotor situado antecedente ao segmento gênico  $L_HV_H$ . Esse acentuador é histoespecífico; ele ativa a transcrição apenas em células B,

não tendo efeito em outros tipos de células. Supostamente, o processo de ativação requer a presença de um fator de transcrição que está presente em linfócitos B mas não em outros tipos de células.

#### PONTOS IMPORTANTES

- Segmentos gênicos adicionais não são transcritos em níveis significativos em células tronco.
- Sua transcrição é ativada pelos rearranjos genômicos que ocorrem durante a diferenciação de linfócitos B.
- Ocorre apenas um rearranjo gênico produtivo de anticorpos por célula.
- A transcrição de genes de cadeia pesada montados é ativada pela transferência de seus promotores para posições perto de um forte acentuador histoespecífico.

## Exercícios Básicos Ilustram a análise genética básica.

1. Qual é a importância do seu sistema imunológico?

**Resposta:** A sua sobrevivência depende de um sistema imunológico funcional. O HIV destrói células que são essenciais ao sistema imunológico, deixando os pacientes com AIDS incapacitados de sobreviver a infecções por patógenos que preocupam pouco ou nada a pessoas com um sistema imunológico funcional.

2. O que são anticorpos e como eles nos protegem de vírus, bactérias e outros patógenos?

**Resposta:** Anticorpos são proteínas produzidas pelo sistema imunológico que se ligam a substâncias exógenas e facilitam sua destruição e remoção do corpo.

3. O que são células T *killer* e como elas nos protegem de vírus, bactérias e outros patógenos?

**Resposta:** Células T *killer* são glóbulos brancos que têm proteínas receptoras do tipo anticorpo em suas superfícies. Esses receptores são usados para reconhecer substâncias exógenas nas superfícies de células infectadas. A detecção de uma substância exógena na superfície de uma célula por uma proteína receptora estimula a célula T a destruir a célula infectada e removê-la do corpo.

4. Seu sistema imunológico é capaz de produzir cerca de um bilhão de anticorpos diferentes para combater infecções. Como o seu genoma pode estocar tanta informação genética?

**Resposta:** Seu genoma não contém um bilhão de genes diferentes codificando anticorpos. Em vez disso, a informação genética que especifica cadeias de anticorpos é estocada em trechos (segmentos gênicos) em seu DNA da linhagem germinativa. Esses trechos são então montados em genes codificantes de cadeias de anticorpos por rearranjos de DNA que ocorrem durante o desenvolvimento de plasmócitos produtores de anticorpos a partir de linfócitos B. Juntando conjuntos de segmentos gênicos em muitas combinações diferentes, pode ser produzida uma grande quantidade de diversidade com uma quantidade relativamente pequena de DNA (veja Figs. 23.10 e 23.11).

5. Se um gene de anticorpo é montado a partir de três segmentos gênicos diferentes, quantas cadeias diferentes de anticorpos podem ser produzidas se o genoma contiver 5, 20 e 200 cópias dos três segmentos gênicos?

**Resposta:** Se cada gene for montado usando uma cópia de cada segmento gênico, 20.000 ( $5 \times 20 \times 200$ ) genes diferentes serão possíveis. Montando genes de anticorpos a partir de segmentos gênicos durante diferenciação celular, pode ser produzida mais diversidade com genomas menores.



## Testar Seus Conhecimentos Integra conceitos e técnicas diferentes.

1. Se o genoma de um determinado primata contém 300  $V_L$ , 5  $J_L$ , 1  $C_L$ , 150  $V_H$ , 5  $J_H$ , 300  $V_H$ , 4  $J_H$ , 50  $D$  e 10 segmentos gênicos  $C_H$  e eles puderem ser unidos em todas as combinações apropriadas possíveis durante os rearranjos genômicos que ocorrem na diferenciação de linfócitos B, quantos anticorpos diferentes este primata poderia produzir?

**Resposta:** Mais de um bilhão de anticorpos diferentes (1,35 bilhões) poderiam ser montados pela junção combinatória dos vários segmentos gênicos e pelas cadeias leves e pesadas diferentes dos anticorpos. Haveria 2.250 cadeias leves diferentes, consistindo em 1.500 cadeias leves kappa (300 segmentos gênicos  $V_L \times 5$  segmentos  $J_L \times 1$  segmento  $C_L$ ) e 750 cadeias leves lambda (150 segmentos gênicos  $V_L \times 5$  segmentos  $J_L - C_L$ ) e 600.000 cadeias pesadas (300 segmentos gênicos  $V_H \times 4$  segmentos  $J_H \times 50$  segmentos  $D \times 10$  segmentos gênicos  $C_H$ ). O número de combinações de cadeias leves e pesadas seria de 2.250 cadeias leves  $\times$  600.000 cadeias pesadas, ou 1.350.000.000 anticorpos diferentes.

2. Em humanos, a ausência de uma enzima chamada purina-nucleosídeo-fosforilase (PNP) resulta em grave deficiência de células T. Desoxiguanosina, o substrato de PNP, acumula-se na ausência da enzima e é especialmente tóxica para linfócitos T. Que efeito você esperaria que essa deficiência de PNP exerça sobre os sistemas imunológicos celular e humoral de indivíduos com o distúrbio?

**Resposta:** Células T *helper* são necessárias para estimular o desenvolvimento tanto de plasmócitos produtores de anticorpos quanto de células T *killer* (veja Fig. 23.7). Considerando que a desoxiguanosina – o substrato da purina-nucleosídeo-fosforilase – é tóxica para linfócitos T, a ausência de atividade de PNP supostamente deverá bloquear tanto as respostas imunológicas humoral quanto celular em uma pessoa com este distúrbio.

## Questões e Problemas Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 23.1. A informação genética que especifica cadeias de anticorpos estocadas em células da linhagem germinativa está no mesmo formato daquela que especifica a maioria dos outros polipeptídeos?
- 23.2. Quantas cadeias polipeptídicas estão presentes em cada molécula de anticorpo? Quantos sítios de ligação de antígenos estão presentes por anticorpo? Quantos anticorpos diferentes são produzidos em cada plasmócito?
- 23.3. Quais as três fontes diferentes de variabilidade de anticorpos?
- 23.4. Anticorpos são de cinco classes diferentes. O que determina a classe de um anticorpo?
- 23.5. Que classe de anticorpo é sempre produzida primeiro em uma célula B em desenvolvimento?
- 23.6. Que classe de imunoglobulina é mais comum no sangue humano?
- 23.7. Que classe de imunoglobulina se liga a receptores em mastócitos e estimula a secreção de histaminas e outras aminas que estão envolvidas em uma reação alérgica?
- 23.8. Considerando o tamanho do genoma de mamíferos, por que não é possível haver um gene separado codificando cada polipeptídeo de anticorpo?
- 23.9. Mudança de classe durante diferenciação do linfócito B ocorre no DNA ou no RNA? Por que mecanismo ela ocorre?
- 23.10. (a) Cite três tipos diferentes de leucócitos que têm papéis importantes na resposta imunológica de vertebrados. (b) Que função cada um destes tipos de células desempenha na resposta imune?
- 23.11. Que importância tem o *locus* de MHC de humanos na prática médica?
- 23.12. Diz-se que os genes do *locus* de MHC de humanos são altamente polimórficos. (a) O que isto significa? (b) Que significado tem isto no campo da medicina?
- 23.13. Quais são as duas principais vias pelas quais complexos anticorpo/antígeno são eliminados?
- 23.14. O que é um epítipo e qual o seu papel no desencadeamento de uma resposta imune?
- 23.15. Onde são produzidas moléculas tais como citocinas, linfocinas e interleucinas e qual é o seu papel na resposta imunológica?
- 23.16. Plasmócitos constituem uma população de células que coletivamente produzem uma ampla gama de anticorpos diferentes. Cada plasmócito produz apenas um único tipo de anticorpo (às vezes anticorpos um pouco diferentes, mas todos específicos para um único antígeno). Após exposição de um humano ou outro vertebrado a um antígeno exógeno tal como a proteína do envoltório de um vírus, o indivíduo inicia rapidamente a produção de grandes quantidades de anticorpos específicos contra o antígeno exógeno ao qual ele foi exposto. Por que mecanismo esta resposta ocorre?



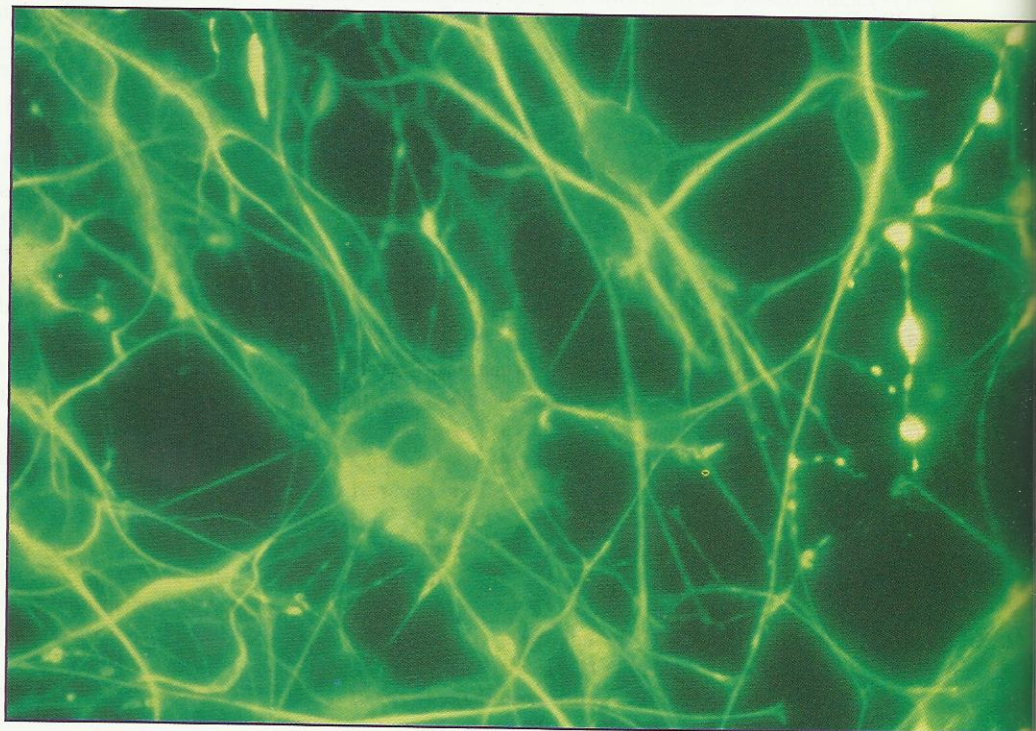
- 23.17. Anticorpos são codificados por famílias de segmentos gênicos que são montados em genes que dirigem sua síntese durante o desenvolvimento dos plasmócitos que os produzem. Existem outras proteínas de mamíferos codificadas por segmentos gênicos que são montados de modo similar durante diferenciação celular?
- 23.18. Na diferenciação de linfócitos B, o que controla eventos de recombinação somática de modo que o segmento do gene  $V$  é unido a um segmento  $J$  e não a outro segmento  $V$  ou diretamente a um segmento  $C$ ?
- 23.19. De que modo as estruturas de anticorpos e os receptores de antígenos de células T são similares?
- 23.20. Por que respostas imunes secundárias são mais rápidas e mais potentes do que respostas imunes primárias?
- 23.21. Infecções microbianas de mamíferos ativam várias respostas diferentes, algumas específicas contra o agente invasor e outras inespecíficas. Quais são as principais respostas imunes inespecíficas? Quais são as principais respostas imunes específicas?
- 23.22. De que modo células de memória diferem da maioria de outras células do sistema imunológico?
- 23.23. Os anticorpos monoclonais demonstraram ser ferramentas muito poderosas para estudos de processos biológicos.
- (a) Como os anticorpos monoclonais são produzidos? (b) Como eles são usados em pesquisa biológica?
- 23.24. Quantas cadeias pesadas de anticorpos diferentes geralmente são produzidas por um determinado plasmócito? Que papel a exclusão alélica tem na produção de anticorpos por plasmócitos?
- 23.25. Células T *helper* são necessárias para montar respostas imunológicas mediadas tanto por anticorpo quanto por células T. Por quê?
- 23.26. Quais são os papéis de fagócitos e proteínas do complemento, respectivamente, nas respostas imunológicas de mamíferos? De que modos as atividades de fagócitos e proteínas do complemento dependem da produção de anticorpos?
- 23.27. O que são doenças auto-imunes e qual o motivo desta denominação?
- 23.28. Se o genoma de um determinado mamífero contém 200 segmentos gênicos  $V_L$ , 4  $J_L$ , 1  $C_L$ , 150  $V_H$ , 6  $J_H$ - $C_H$ , 200  $V_H$ , 4  $J_H$ , 100  $D$  e 8  $C_H$  e eles puderem ser unidos em todas as combinações apropriadas possíveis durante os rearranjos do genoma que ocorrem na diferenciação de linfócitos B, quantos anticorpos diferentes este mamífero poderia produzir?
- 23.29. Proponha um mecanismo pelo qual ocorre uma doença auto-imune tal como diabetes juvenil.



# A Base Genética do Câncer

## Tópicos

- Câncer: Uma Doença Genética
- Oncogenes
- Genes Supressores Tumorais
- Vias Genéticas para o Câncer



Micrografia de luz fluorescente de uma cultura de células humanas de tumor cerebral. As células são interconectadas por axônios fibrosos e dendritos.

## Um Assassino Comum

Lawrence Taylor, um professor de biologia em uma grande universidade americana, teve dificuldade em concluir uma de suas palestras. Seus pensamentos eram complexos e desfocados, e ele se sentia cansado e desorientado. Após dispensar os estudantes, Lawrence procurou um médico. O diagnóstico preliminar foi que o estado mental e as capacidades motoras de Lawrence haviam sido prejudicados por um pequeno derrame, resultante de rompimento de um pequeno vaso sanguíneo no cérebro. Entretanto, o exame posterior revelou um problema mais grave. O cérebro de Lawrence havia desenvolvido um tumor, provavelmente derivado de outro tumor situado em um de seus rins. Estas duas massas de tecido canceroso tinham crescido a ponto de começar



a interferir em funções biológicas importantes. Lawrence submeteu-se a uma cirurgia para remover o tumor renal e recebeu terapia de radiação para destruir o tumor cerebral. Infelizmente, o crescimento do tumor cerebral não pôde ser detido, e, 18 meses após dar sua última aula, Lawrence morreu.

## CÂNCER: UMA DOENÇA GENÉTICA

Mutações em genes que controlam o crescimento e a divisão celular são responsáveis pelo câncer.

Tumores cancerosos matam centenas de milhares de americanos a cada ano. O que causa a formação de tumores e o que faz com que alguns deles disseminem? Por que alguns tipos de tumores tendem a ser encontrados em famílias? A tendência em desenvolver o câncer é hereditária? Fatores ambientais contribuem para o desenvolvimento do câncer? Nos últimos anos, estas e outras perguntas estimularam uma enorme quantidade de pesquisas sobre a biologia básica do câncer. Embora muitos detalhes ainda não estejam claros, o achado fundamental é que cânceres resultam de disfunções genéticas. Em alguns casos, essas disfunções podem ser ativadas ou exacerbadas por fatores ambientais tais como dieta, exposição excessiva ao sol ou poluentes químicos. Os cânceres surgem quando genes críticos são mutados. Essas mutações podem causar mau funcionamento de processos bioquímicos e resultar em uma proliferação desregulada de células. Sem regulação, as células cancerosas dividem-se incessantemente, acumulando-se umas sobre as outras para formar tumores. Quando células se destacam de um tumor e invadem os tecidos vizinhos, o tumor é maligno. Quando as células não invadem os tecidos vizinhos, o tumor é benigno. Tumores malignos podem disseminar-se para outros locais no corpo, formando tumores secundários. Este processo é denominado **metástase**, das palavras em grego que significam “estado alterado”. Tanto em tumores benignos quanto em malignos, algo deu errado com os sistemas que

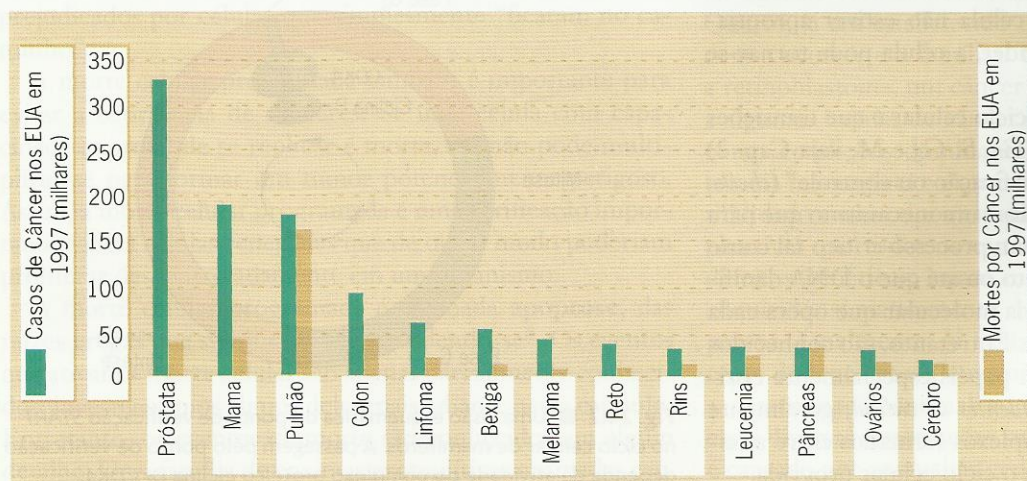
controlam a divisão celular. Pesquisadores estabeleceram firmemente que esta perda de controle se deve a defeitos genéticos subjacentes.

## As Muitas Formas de Câncer

O câncer não é uma doença única, mas, em vez disso, um grupo de doenças. Os cânceres podem originar-se em muitos tecidos diferentes do corpo. Alguns crescem agressivamente; outros, mais lentamente. Alguns tipos de câncer podem ser detidos por tratamento médico apropriado; outros, não. A Fig. 24.1 mostra as frequências de novos casos de tipos diferentes de câncer nos EUA, bem como o número de fatalidades atribuídas a cada tipo. O tipo mais comum de câncer envolve a glândula prostática em homens. O câncer de mama em mulheres e o câncer de pulmão em ambos os sexos também são comuns. Em alguns casos, uma predisposição a desenvolver câncer é herdada. Entretanto, a maioria dos casos de câncer decorre do acúmulo de mutações espontâneas nos tecidos somáticos.

Os tipos mais prevalentes de câncer são derivados de populações de células que se dividem ativamente – por exemplo, células epiteliais nos intestinos, nos pulmões ou na glândula prostática. As formas mais raras de câncer desenvolvem-se de populações celulares que tipicamente não se dividem – por exemplo, células diferenciadas musculares ou nervosas.

Embora a taxa de morte por câncer ainda seja alta, foi feito um enorme progresso para detectar e tratar tipos diferentes de câncer. As técnicas de genética molecular capacitaram cientistas a caracterizar cânceres de modos que antes não eram possíveis e lhes permitiram criar novas estratégias



**Fig. 24.1** ■ Número estimado de novos casos e mortes de tipos específicos de câncer nos EUA em 1997.



para terapia do câncer. Há poucas dúvidas de que um grande investimento em pesquisa básica do câncer está sendo compensador.

Células cancerosas podem ser obtidas para estudos experimentais removendo-se tecido de um tumor e o dissociando em suas células constituintes. Com nutrientes apropriados, tais células tumorais dissociadas podem ser cultivadas *in vitro*, às vezes indefinidamente. Células cancerosas também podem ser derivadas de culturas de células normais tratando-se as células com agentes que induzem o estado canceroso. Radiação, substâncias químicas mutagênicas e alguns tipos de vírus podem transformar irreversivelmente células normais em células cancerosas. Os agentes que causam este tipo de transformação são chamados de **carcinógenos**.

A característica marcante de todas as células cancerosas é que seu crescimento é desregulado. Quando células normais são cultivadas *in vitro*, elas formam uma única camada – uma monocamada – na superfície do meio de cultura. Células cancerosas, em contraste, crescem umas por cima das outras, acumulando-se na superfície do meio de cultura para formar massas. Esse acúmulo desregulado ocorre porque células cancerosas não respondem aos sinais químicos que inibem a divisão celular e porque não podem formar associações estáveis com suas vizinhas.

As anomalias externas que são aparentes em uma cultura de células cancerosas estão correlacionadas com profundas anomalias intracelulares. Células cancerosas em geral têm um citoesqueleto desorganizado; elas podem sintetizar proteínas incomuns e apresentá-las em suas superfícies e frequentemente têm números cromossômicos anormais, isto é, elas são aneuploides.

## Câncer e o Ciclo Celular

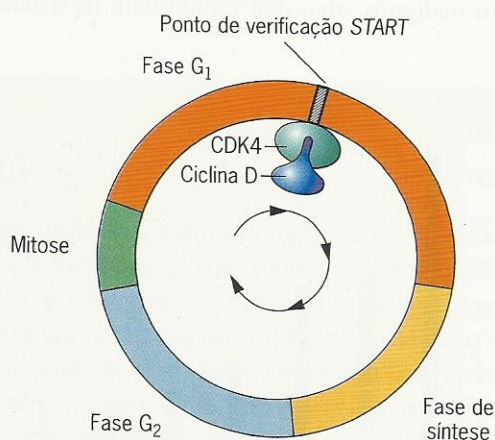
O ciclo celular consiste em períodos de crescimento, síntese de DNA e divisão. O tamanho desse ciclo e a duração de cada um de seus componentes são controlados por sinais químicos externos e internos. A transição de cada fase do ciclo requer a integração de sinais químicos específicos e respostas precisas a estes sinais. Se os sinais forem incorretamente percebidos ou se a célula não estiver apropriadamente preparada para responder, a célula pode tornar-se cancerosa.

A visão atual do controle do ciclo celular é que transições entre fases diferentes do ciclo ( $G_1$ , S,  $G_2$  e M; veja Cap. 2) são reguladas em “pontos de verificação ou controle” (*check-points*). Um **ponto de verificação** é um mecanismo que pára a progressão pelo ciclo até que um processo crítico tal como a síntese de DNA esteja completo ou até que o DNA danificado seja reparado. A maquinaria molecular que opera cada ponto de verificação é complexa. Entretanto, são conhecidos dois tipos de proteínas que têm papéis especialmente críticos: as *ciclínas* e as *cinases dependentes de ciclina*, geralmente abreviadas como CDK. Os complexos formados entre as ciclínas e as CDK fazem com que o ciclo celular progrida.

As CDK são componentes cataliticamente ativos do mecanismo de ciclo celular. Tais proteínas regulam as atividades de outras proteínas transferindo grupos fosfato para elas. Entretanto, a atividade de fosforilação das CDK depende da presença de ciclínas. As ciclínas capacitam as CDK a desempenhar a sua função formando complexos ciclina/CDK. Quando as ciclínas estão ausentes, estes complexos não podem ser formados, e as CDK são inativas. O ciclo celular requer, portanto, a formação e a degradação alternadas de complexos ciclina/CDK.

Um dos mais importantes pontos de verificação do ciclo celular, chamado **START**, está no meio de  $G_1$  (Fig. 24.2). A célula recebe tanto sinais externos quanto internos neste ponto de verificação para determinar quando é apropriado passar para a fase S. Esse ponto de verificação é regulado por ciclínas tipo D em conjunto com CDK<sub>4</sub>. Se uma célula é ativada a ir além do ponto de verificação **START** pelo complexo ciclina D/CDK<sub>4</sub>, ela entra em outra rodada de replicação do DNA. Proteínas inibitórias com a capacidade de perceber problemas na fase tardia de  $G_1$ , tais como baixos níveis de nutrientes ou dano ao DNA, podem colocar um freio no complexo ciclina/CDK e impedir que a célula entre na fase S. Na ausência de tais problemas, o complexo ciclina D/CDK<sub>4</sub> ativa a célula para ir do final da fase  $G_1$  para a fase S, iniciando assim a replicação do DNA que é um prelúdio da divisão celular.

Em células tumorais, os pontos de verificação do ciclo celular estão tipicamente desregulados. Essa desregulação se deve a defeitos genéticos na maquinaria que alternadamente aumenta e diminui a abundância dos complexos ciclina/CDK. Por exemplo, os genes que codificam as ciclínas ou as CDK podem estar mutados, ou os genes que codificam as proteínas que respondem a complexos específicos ciclina/CDK ou que regulam a abundância destes complexos podem estar mutados. Muitos tipos diferentes de defeitos genéticos podem desregular o ciclo celular, com a consequência final de que as células podem tornar-se cancerosas.



**Fig. 24.2** ■ Uma visão esquemática do ponto de verificação **START** no ciclo celular de mamíferos. A passagem pelo ponto de verificação depende da atividade do complexo protéico ciclina D/CDK<sub>4</sub>.



Células em que o ponto de verificação *START* é disfuncional são especialmente propensas a tornar-se cancerosas. O ponto de verificação *START* controla a entrada na fase S do ciclo celular. Se o DNA dentro de uma célula foi danificado, é importante que a entrada na fase S seja retardada para permitir reparo do DNA danificado. De outro modo, o DNA danificado será replicado e transmitido para todas as descendentes da célula. Células normais são programadas para fazer uma pausa no ponto de verificação *START* para garantir que o reparo seja completado antes que comece a replicação do DNA. Em contraste, células em que o ponto de verificação *START* é disfuncional movem-se para a fase S sem reparar seu DNA danificado. Durante uma série de ciclos celulares, mutações que resultam da replicação de DNA não reparado podem acumular-se e causar mais desregulação do ciclo celular. Um clone de células com um ponto de verificação *START* disfuncional pode, portanto, tornar-se agressivamente canceroso.

### Câncer e Morte Celular Programada

Todo câncer envolve o acúmulo de células indesejadas. Em muitos animais, células supérfluas podem ser descartadas por mecanismos que são programados pelas próprias células. Essa morte celular programada foi originalmente descoberta em estudos com o nematódeo *Caenorhabditis elegans*. Este pequeno nematelminto perde algumas das células que se acumulam durante os cerca de 10 ciclos de divisão que ocorrem no período de seu desenvolvimento a partir de um zigoto. Análises genéticas de Robert Horvitz e seus colaboradores demonstrou que a perda destas células não ocorre em algumas linhagens mutantes de *C. elegans*. Assim, a morte celular é parte do programa normal de desenvolvimento neste animal – e em outros também. Sabemos, por exemplo, que, durante o desenvolvimento das mãos e dos pés de muitos vertebrados, as células que ficam entre os dedos em desenvolvimento devem morrer. Caso contrário, os dedos continuarão fundidos. A morte celular programada é, portanto, um fenômeno fundamental e amplo entre animais. Sem ela, a formação e o funcionamento dos órgãos seriam prejudicados por células que simplesmente “ficaram no caminho”.

A morte celular programada também é importante para evitar a ocorrência de cânceres. Se uma célula com capacidade anormal de se replicar é morta, ela não pode multiplicar-se para formar um tumor potencialmente perigoso. Assim, a morte celular programada é uma verificação importante contra células renegadas que de outro modo poderiam proliferar descontroladamente em um organismo.

A morte celular programada é chamada **apoptose**, das raízes gregas que significam “desprendendo-se”. Os eventos que ativam a morte celular são apenas parcialmente conhecidos; investigaremos alguns deles mais adiante neste capítulo. Entretanto, os eventos de morte são conhecidos em algum detalhe. Uma família de enzimas proteolíticas chamadas cas-

pases desempenha um papel crucial no fenômeno de morte celular. As caspases removem pequenas partes de outras proteínas cortando ligações peptídicas. Por meio deste aparato enzimático, as proteínas alvos são inativadas. As caspases atacam muitos tipos diferentes de proteínas, incluindo as lâminas, que constituem o revestimento interno do envoltório nuclear, e vários componentes do citoesqueleto. O impacto coletivo dessa clivagem proteolítica é que células nas quais ela ocorre perdem sua integridade; sua cromatina torna-se fragmentada, formam-se bolhas de citoplasma em sua superfície e elas começam a encolher. Células que sofrem este tipo de desintegração geralmente são englobadas por fagócitos, que são células removedoras do sistema imunológico, e são então destruídas. Se o mecanismo apoptótico foi prejudicado ou inativado, uma célula que de outro modo seria morta pode sobreviver e proliferar. Tal célula tem o potencial de formar um clone que pode tornar-se canceroso se obtiver a capacidade de se dividir descontroladamente.

### Uma Base Genética para o Câncer

Os grandes avanços recentes na compreensão do câncer foram proporcionados pela aplicação de técnicas de genética molecular. Entretanto, antes destas técnicas estarem disponíveis para pesquisadores, houve forte evidência de que as causas subjacentes do câncer são genéticas. Primeiro, sabia-se que o estado canceroso é herdado clonalmente. Quando células cancerosas são mantidas em cultura, suas descendentes são todas cancerosas. A condição cancerosa é portanto transmitida de cada célula para suas filhas na época da divisão, um fenômeno revelador de que o câncer tem uma base genética. Segundo, sabia-se que alguns tipos de vírus podem induzir a formação de tumores em animais experimentais. A indução de câncer por vírus significa que as proteínas codificadas por genes virais estão envolvidas na produção do estado canceroso. Terceiro, sabe-se que o câncer pode ser induzido por agentes capazes de causar mutações. Foi demonstrado que substâncias mutagênicas e radiação ionizante induzem tumores em animais experimentais. Além disso, um grande volume de dados epidemiológicos implicou estes agentes como causadores de câncer em seres humanos. Quarto, já era estabelecido que alguns tipos de câncer tendem a ocorrer em famílias. Em particular, a suscetibilidade a retinoblastoma, um câncer raro de olho, e a predisposição a algumas formas de câncer de cólon parecem ser herdadas como condições dominantes simples, embora com penetração incompleta e expressividade variável. Como a suscetibilidade a estes tipos especiais de câncer é herdada, parece plausível que todos os cânceres possam estar baseados em defeitos genéticos, sejam mutações herdadas ou mutações adquiridas somaticamente durante o tempo de vida de uma pessoa. Finalmente, já sabíamos que alguns tipos de cânceres de leucócitos (leucemias e linfomas) estão associados a determinadas anomalias cromossômicas. Coletivamente, essas diversas observações sugerem fortemente que o câncer é causado por anomalias genéticas.



Nos anos 1980, quando técnicas de genética molecular foram primeiro usadas para estudar células cancerosas, pesquisadores descobriram que o estado canceroso é, de fato, associável a defeitos genéticos específicos. Tipicamente, entretanto, não apenas um mas vários desses defeitos são necessários para converter uma célula normal em uma célula cancerosa. Pesquisadores de câncer identificaram duas grandes classes de genes que, quando mutadas, podem contribuir para o desenvolvimento de um estado canceroso. Em uma dessas classes, genes mutantes promovem ativamente a divisão celular; na outra classe, genes mutantes não reprimem a divisão celular. Os genes na primeira classe são chamados **oncogenes**, da raiz grega que significa "tumor". Os genes na segunda classe são chamados de **genes supressores tumorais**. Nas seções seguintes, discutiremos a descoberta, as características e o significado de cada uma destas classes de genes relacionados ao câncer.

### PONTOS IMPORTANTES

- O câncer é um grupo de doenças em que o ciclo celular de crescimento e divisão está desregulado.
- Cânceres podem desenvolver-se caso o mecanismo de morte celular programada (apoptose) esteja prejudicado.
- Cânceres são oriundos da ocorrência de mutações em genes cujos produtos protéicos estão envolvidos no controle do ciclo celular.

## ONCOGENES

Muitos cânceres envolvem a hiperexpressão de determinados genes ou a atividade anormal de seus produtos protéicos mutantes.



Os oncogenes constituem um grupo diverso de genes cujos produtos têm papéis importantes na regulação de atividades bioquímicas dentro das células, incluindo aquelas atividades relaciona-

das a divisão celular. Esses genes foram primeiro descobertos nos genomas de vírus com RNA que são capazes de induzir tumores em vertebrados hospedeiros. Mais tarde, as contrapartes celulares desses oncogenes virais foram descobertas em muitos organismos diferentes, variando de *Drosophila* a seres humanos.

## Retrovírus Indutores de Tumor e Oncogenes Virais

*Insights* fundamentais sobre a base genética do câncer resultaram do estudo de vírus indutores de tumor. Muitos desses vírus têm um genoma composto de RNA em lugar de DNA. Após entrar em uma célula, o RNA viral é usado como molde para sintetizar DNA complementar, que então é inserido em uma ou mais posições nos cromossomos das células. A síntese de DNA a partir de RNA é catalisada pela enzima

viral transcriptase reversa. Essa reversão do fluxo normal da informação genética do DNA para o RNA motivou biólogos a chamar tais patógenos de **retrovírus** (veja Cap. 18).

O primeiro vírus indutor de tumor foi descoberto em 1910 por Peyton Rous. Ele causa um tipo especial de tumor, ou sarcoma, no tecido conjuntivo de galinhas e foi então chamado de vírus do sarcoma de Rous. Pesquisas modernas mostraram que o genoma de RNA deste retrovírus contém quatro genes: *gag*, que codifica a proteína do capsídeo do vírion; *pol*, que codifica a transcriptase reversa; *env*, que codifica uma proteína do envoltório viral; e *v-src*, que codifica uma cinase protéica (proteína cinase) que se insere nas membranas plasmáticas das células infectadas. A característica distintiva de uma cinase é que ela pode fosforilar outras proteínas. Destes quatro genes, apenas o gene *v-src* é responsável pela capacidade do vírus de formar tumores. Um vírus em que o gene *v-src* foi deletado é infeccioso mas incapaz de induzir tumores. Genes tais como o *v-src* que causam câncer são denominados oncogenes.

Estudos com outros retrovírus indutores de tumor descobriram pelo menos 20 oncogenes virais diferentes, em geral incidados como *v-onc* (Quadro 24.1). Alguns deles estão relacionados a genes celulares que codificam fatores de crescimento. Por exemplo, *v-sis*, um oncogene do vírus do sarcoma símio, codifica uma versão do fator de crescimento derivado de plaquetas (PDGF). PDGF normalmente é produzido por plaquetas para promover cicatrização de feridas, o que ocorre por estimulação do crescimento de células no local da ferida. O vírus do sarcoma de símios que leva o gene *v-sis* induz tumores em macacos. Eles também transformam células cultivadas para um estado canceroso, supostamente produzindo grandes quantidades da versão *v-sis* de PDGF, que então causa crescimento celular descontrolado.

Outros oncogenes virais codificam proteínas que são similares a fatores de crescimento e receptores hormonais. Por exemplo, o gene *v-erbB* do vírus de eritroblastose aviária codifica uma proteína muito similar ao receptor celular do fator de crescimento epidérmico (EGF) e o gene *v-fms* do vírus do sarcoma felino codifica uma proteína muito similar ao receptor do fator de crescimento celular CSF-1 (fator 1 estimulante de colônia). Ambos estes receptores de fatores de crescimento são proteínas transmembranares com um domínio de ligação ao fator de crescimento do lado externo da célula e um domínio de cinase protéica no lado interno. Este domínio no interior permite que a proteína fosforile alguns aminoácidos, geralmente tirosina, em outras proteínas que interagem com ela.

Muitos oncogenes virais, incluindo *v-src*, codificam tirosina-cinases não inseridas na membrana plasmática. Em vez disso, tais proteínas estão situadas na face interna da membrana plasmática, onde efetuam sua função de fosforilação. Os vários oncogenes *v-ras* codificam proteínas que ligam GTP, como as proteínas G celulares, que têm um papel importante na regulação do nível de AMP cíclico.

Outro grupo de oncogenes virais codifica proteínas que aparentemente funcionam como fatores de transcrição. Eles incluem os genes *v-jun*, *v-fos*, *v-erbA* e *v-myc*, cada um levado



QUADRO 24.1

## Oncogenes Retrovirais

| Oncogene         | Vírus                                   | Espécie de Hospedeiro | Função do Produto Gênico                                                   |
|------------------|-----------------------------------------|-----------------------|----------------------------------------------------------------------------|
| <i>abl</i>       | Vírus da leucemia murina Abelson        | Camundongo            | Cinase protéica específica de tirosina                                     |
| <i>erbA</i>      | Vírus da eritroblastose aviária         | Galinha               | Análogo de receptor hormonal de tireóide                                   |
| <i>erbB</i>      | Vírus da eritroblastose aviária         | Galinha               | Versão truncada de receptor de fator de crescimento epidérmico (EGF)       |
| <i>fes</i>       | Vírus do sarcoma felino ST              | Gato                  | Cinase protéica específica de tirosina                                     |
| <i>fgr</i>       | Vírus do sarcoma felino Gardner-Rasheed | Gato                  | Cinase protéica específica de tirosina                                     |
| <i>fms</i>       | Vírus do sarcoma felino McDonough       | Gato                  | Análogo de receptor de fator de crescimento estimulante de colônia (CSF-1) |
| <i>fos</i>       | Vírus do osteossarcoma FJB              | Camundongo            | Proteína ativadora transcricional                                          |
| <i>fps</i>       | Vírus do sarcoma Fuginami               | Galinha               | Cinase protéica específica de tirosina                                     |
| <i>jun</i>       | Vírus 17 do sarcoma aviário             | Galinha               | Proteína ativadora transcricional                                          |
| <i>mil (mht)</i> | Vírus MH2                               | Galinha               | Cinase protéica de serina/treonina                                         |
| <i>mos</i>       | Vírus do sarcoma Moloney                | Camundongo            | Cinase protéica de serina/treonina                                         |
| <i>myb</i>       | Vírus da mieloblastose aviária          | Galinha               | Fator de transcrição                                                       |
| <i>myc</i>       | Vírus da mielocitomatose MC29           | Galinha               | Fator de transcrição                                                       |
| <i>raf</i>       | Vírus do sarcoma murino 3611            | Camundongo            | Cinase protéica de serina/treonina                                         |
| <i>H-ras</i>     | Vírus do sarcoma murino Harvey          | Rato                  | Proteína de ligação de GTP                                                 |
| <i>K-ras</i>     | Vírus do sarcoma murino Kirsten         | Rato                  | Proteína de ligação de GTP                                                 |
| <i>rel</i>       | Vírus da reticuloendoteliose            | Peru                  | Fator de transcrição                                                       |
| <i>ros</i>       | Vírus do sarcoma aviário UR11           | Galinha               | Cinase protéica específica de tirosina                                     |
| <i>sis</i>       | Vírus do sarcoma de símios              | Macaco                | Análogo de fator de crescimento derivado de plaquetas (PDGF)               |
| <i>src</i>       | Vírus do sarcoma de Rous                | Galinha               | Cinase protéica específica de tirosina                                     |
| <i>yes</i>       | Vírus do sarcoma Y73                    | Galinha               | Cinase protéica específica de tirosina                                     |

por um retrovírus diferente. Essas proteínas são homólogas a proteínas celulares que se ligam ao DNA e regulam transcrição.

Cada tipo de oncogene viral parece, portanto, codificar uma proteína que pode teoricamente desempenhar um papel em regular a expressão de genes celulares, incluindo aqueles envolvidos no processo de crescimento e divisão. Algumas dessas proteínas podem agir como sinais para estimular alguns tipos de atividade celular; outras podem agir como receptores para captar esses sinais ou como agentes intracelulares para levá-las da membrana celular para o núcleo. Uma outra categoria de proteínas oncogênicas virais pode agir como fatores de transcrição para estimular expressão gênica.

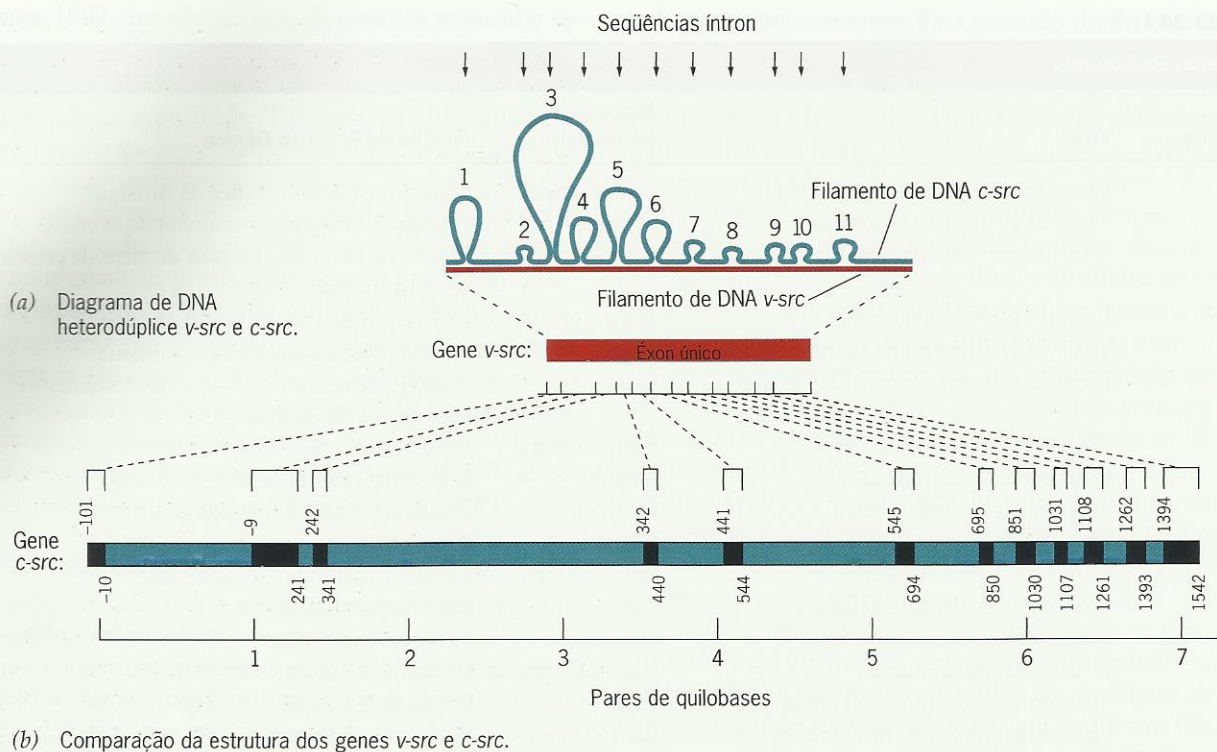
### Homólogos Celulares de Oncogenes Virais: Os Protoncogenes

As proteínas codificadas por oncogenes virais são similares a proteínas celulares com importantes funções reguladoras. Muitas dessas proteínas celulares foram identificadas mediante isolamento do homólogo celular do oncogene viral. Por exemplo, o homólogo celular do gene *v-src* foi obtido

triando-se uma biblioteca de DNA genômico feita de células de galinha não infectadas. Para essa triagem, o gene *v-src* foi usado como uma sonda de hibridização para detectar clones de DNA recombinante que poderiam fazer pares de bases com ela. A análise desses clones estabeleceu que células de galinha contêm um gene que é similar ao *v-src*, isto é, relacionado a ele no sentido evolutivo. Entretanto, este gene não está associado a um provírus integrado de sarcoma, e ele difere do gene *v-src* em um aspecto muito importante: ele contém íntrons (Fig. 24.3). Existem, de fato, 11 íntrons no homólogo de galinha do *v-src*, comparados a zero no próprio gene *v-src*. Esta incrível descoberta sugeriu que talvez *v-src* evoluiu de um gene celular normal e que, concomitantemente, perdeu seus íntrons.

Os homólogos celulares dos oncogenes virais são chamados de **protoncogenes** ou, às vezes, **oncogenes celulares normais**, indicados por *c-onc*. O homólogo celular de *v-src* é portanto *c-src*. As seqüências codificantes destes dois genes são muito similares, diferindo apenas em 18 nucleotídeos; *v-src* codifica uma proteína com 526 aminoácidos e *c-src* codifica uma proteína com 533 aminoácidos. Usando genes *v-onc* como sondas, outros genes *c-onc* foram isolados de muitos organismos diferentes, incluindo seres humanos. Como regra, esses oncogenes celulares mostram uma considerável





**Fig. 24.3 ■ Estruturas dos genes v-src e c-src.** (a) Diagrama do heteroduplex de DNA previsto formado por hibridização de um filamento do gene c-src (em cima) e um filamento parcialmente complementar do gene v-src (embaixo). Os íntrons (numerados de 1 a 11) formam alças unifilamentares. (b) Comparação esquemática destes dois genes, com éxons mostrados em preto. O sistema coordenado para éxons no gene c-src é baseado no primeiro nucleotídeo na sequência codificante (posição 1). O primeiro exon (posição -101 a -10) está no líder 5' do gene.

conservação em estrutura. *Drosophila*, por exemplo, possuem homólogos muito similares dos oncogenes celulares de vertebrados *c-abl*, *c-erbB*, *c-fps*, *c-raf*, *c-ras* e *c-myb*. A similaridade de oncogenes de espécies diferentes sugere fortemente que as proteínas que eles codificam estão envolvidas em importantes funções celulares.

Por que *c-oncs* têm íntrons ao passo que *v-oncs* não? A resposta mais plausível é que *v-oncs* foram derivados de *c-oncs* pela inserção de um mRNA de *c-onc* totalmente processado no genoma de um retrovírus. Um vírion que contenha tal molécula recombinante seria capaz de transduzir o gene *c-onc* sempre que infectar outra célula. Durante infecção, o RNA recombinante seria reversamente transcrito em DNA e então integrado aos cromossomos da célula. O que poderia ser de maior valor para um vírus do que ter um novo gene que estimula crescimento aumentado de seu hospedeiro, enquanto seu genoma integrado se replica junto com o da célula?

Em muitos casos, a aquisição de um oncogene por um retrovírus foi acompanhada da perda de algum material genético viral. Como esse material perdido é necessário para replicação viral, tais vírus oncogênicos são capazes de se reproduzir apenas se um vírus auxiliar estiver presente. Quanto a isto, eles se assemelham aos bacteriófagos transdutores que discutimos no Cap. 8.

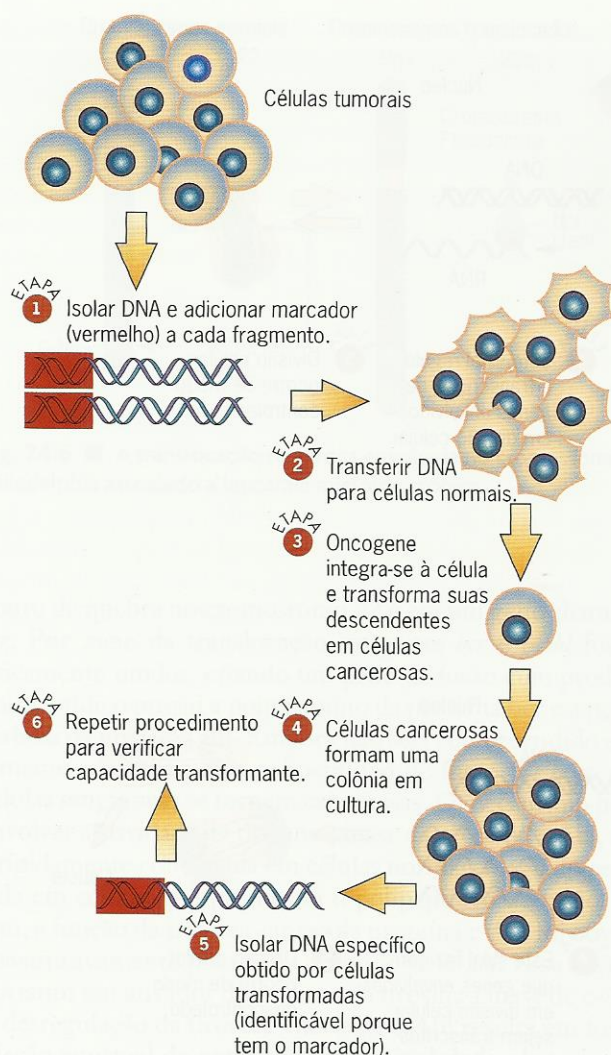
Por que *v-oncs* induzem tumores, enquanto *c-oncs* normais não? Em alguns casos, parece que o oncogene viral produz muito mais proteína do que sua contraparte celular, talvez porque ele tenha sido transcricionalmente ativado por acentuadores inseridos no genoma viral. Em células tumorais de galinha, por exemplo, o gene *v-src* produz 100 vezes mais tirosina-cinase do que o gene *c-src*. Este enorme excesso de cinase evidentemente perturba os delicados mecanismos de sinalização que controlam a divisão celular, causando crescimento desregulado. Outros genes *v-onc* podem induzir tumores expressando suas proteínas em momentos impróprios ou expressando formas alteradas, isto é, mutadas, destas proteínas.

### Oncogenes Celulares Mutantes e Câncer

Os produtos dos *c-oncs* exercem papéis centrais em regular atividades celulares. Conseqüentemente, uma mutação em um desses genes pode comprometer o equilíbrio bioquímico dentro de uma célula e colocá-la em uma via para tornar-se cancerosa. Estudos de muitos tipos diferentes de câncer humano demonstraram que oncogenes celulares mutantes estão associados ao desenvolvimento de um estado canceroso.

A primeira evidência que associa câncer a um *c-onc* mutante resultou do estudo de um câncer humano de bexiga.





**Fig. 24.4** ■ O teste de transfecção para identificar sequências de DNA capazes de transformar células normais em células cancerosas.

A mutação responsável por esse câncer de bexiga foi isolada por Robert Weinberg e colaboradores que usaram um teste de transfecção (Fig. 24.4). O DNA foi extraído de um tecido canceroso e fragmentado em pequenos pedaços. Então, cada um desses pedaços foi unido a um segmento de DNA bacteriano, que serviu como marcador molecular. Os fragmentos de DNA marcados foram então introduzidos, ou transfectados, para células mantidas em cultura para determinar se algumas delas podiam transformar as células em um estado canceroso. Esse estado pode ser reconhecido pela tendência das células cancerosas em formar pequenos grupos, ou focos, quando cultivadas em placas de ágar. O DNA de tais células foi extraído e triado para comprovar se levava o marcador molecular que estava ligado aos fragmentos transfectantes originais. Caso sim, esse DNA era testado novamente quanto a sua capacidade de induzir o estado canceroso. Após vários testes, a pesquisa de Weinberg identificou um fragmento de DNA do câncer de bexiga original que reprodutivamente

transformou células cultivadas em células cancerosas. Esse fragmento levava um alelo do oncogene *c-H-ras*, um homólogo de um oncogene na linhagem Harvey do vírus do sarcoma de rato. A análise da sequência de DNA mostrou subsequentemente que um nucleotídeo no códon 12 deste alelo tinha sido mutado, com substituição de uma glicina por uma valina normalmente encontrada nesta posição na proteína *c-H-ras*.

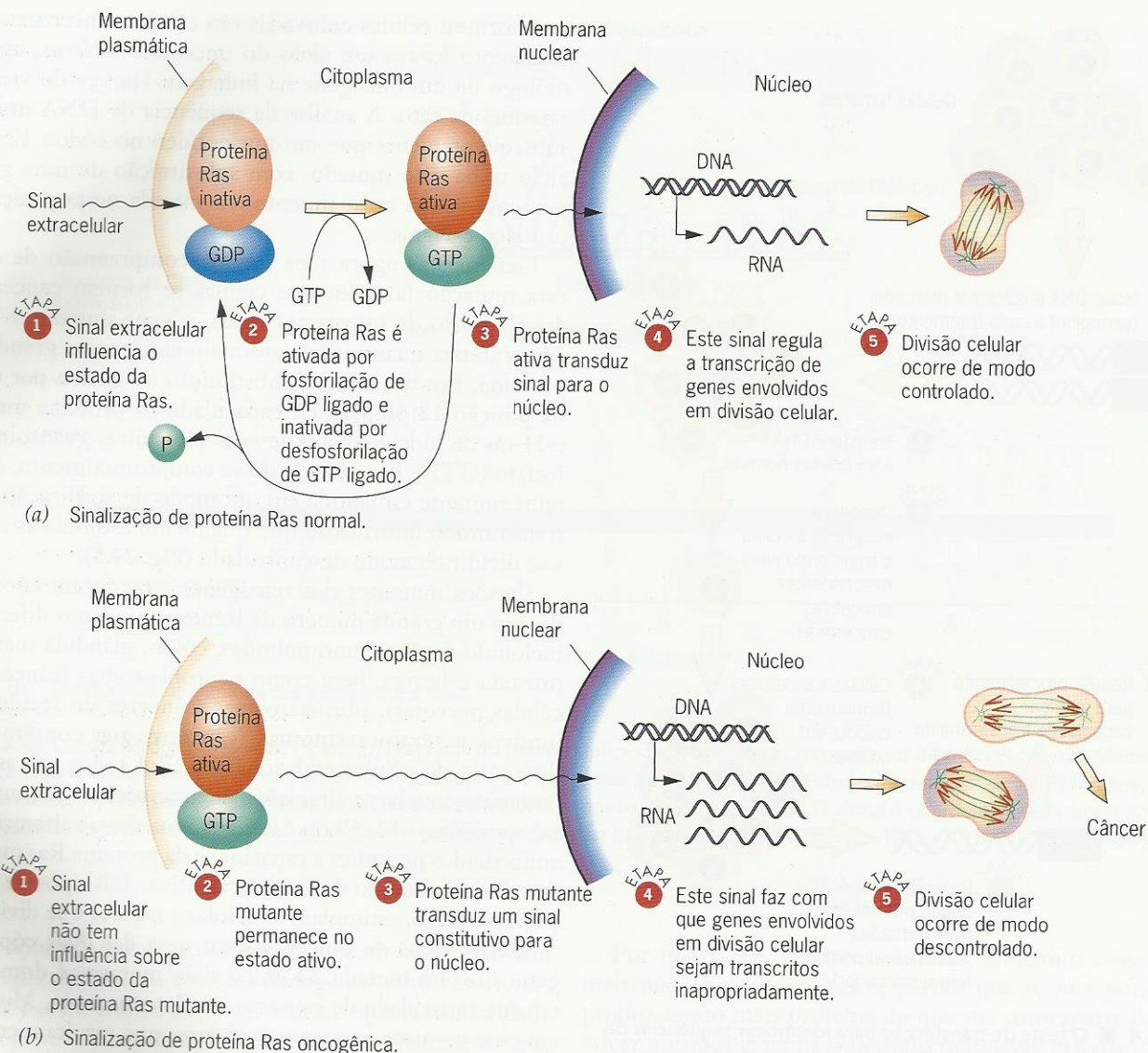
Geneticistas agora têm alguma compreensão de como esta mutação faz com que células se tornem cancerosas. Ao contrário de oncogenes virais, o gene mutante *c-H-ras* não sintetiza quantidades anormalmente muito grandes de proteína. Em vez disso, a substituição de glicina por valina na posição 12 prejudica a capacidade da proteína mutante *c-H-ras* de hidrolisar um de seus substratos, guanossina-trifosfato (GTP). Em virtude desse comprometimento, a proteína mutante é mantida em um modo de sinalização ativa, transmitindo informação que finalmente estimula as células a se dividir de modo descontrolado (Fig. 24.5).

Versões mutantes dos oncogenes *c-ras* foram encontradas em um grande número de tumores humanos diferentes, incluindo órgãos como pulmões, cólon, glândula mamária, próstata e bexiga, bem como neuroblastomas (cânceres de células nervosas), fibrossarcomas (cânceres do tecido conjuntivo) e teratocarcinomas (cânceres que contêm tipos diferentes de células embrionárias). Em todos os casos, as mutações envolvem alterações de aminoácidos em uma das três posições – 12, 59 ou 61. Cada uma dessas alterações de aminoácidos prejudica a capacidade da proteína Ras mutante de sair de seu modo de sinalização ativa. Tais tipos de mutações, portanto, estimulam as células a crescer e se dividir.

Nestes tipos de câncer, apenas uma das duas cópias do gene *c-ras* foi mutada. O único alelo mutante é dominante em sua capacidade de causar o estado canceroso. Mutações em *c-ras* e outros oncogenes celulares que resultam em câncer deste modo são portanto **ativadores dominantes** de crescimento celular descontrolado.

Mutações ativadoras dominantes em oncogenes celulares raramente são herdadas pela linhagem germinativa; em vez disso, a maioria delas ocorre espontaneamente na linhagem somática durante o curso de divisões celulares. Como o número de divisões celulares na vida humana é muito grande, mais de  $10^6$  milhares de mutações potencialmente oncogênicas deverão ocorrer, e, se cada uma funcionar como um ativador dominante de crescimento celular descontrolado, o desenvolvimento de um tumor seria inevitável. Entretanto, muitas pessoas levam vidas longas sem desenvolver tumores. A explicação deste paradoxo é que cada mutação oncogênica individual é, por si só, raramente capaz de induzir um estado canceroso. Em seres humanos, as células aparentemente têm modos de se ajustar aos efeitos deletérios de um único oncogene mutante. Entretanto, quando vários genes diferentes reguladores do crescimento dentro de uma célula foram alterados por mutações, a célula não pode compensar seus efeitos individuais, seu crescimento torna-se desregulado e sobrevém câncer. Em muitos tumores, pelo menos uma des-





**Fig. 24.5** ■ Sinalização da proteína Ras e câncer. (a) O produto protéico normal do gene *ras* alterna entre os estados inativo e ativo, dependendo de sua ligação ser a GDP ou a GTP. Sinais extracelulares tais como fatores de crescimento estimulam a conversão de Ras inativo em Ras ativo. Através de Ras ativo, esses sinais são transmitidos para outras proteínas e finalmente para o núcleo, onde induzem a expressão de genes envolvidos em divisão celular. Como esta sinalização é intermitente e regulada, a divisão celular ocorre de modo controlado. (b) Proteínas Ras mutantes existem principalmente no estado ativo. Essas proteínas transmitem seus sinais mais ou menos constantemente, resultando em divisão celular descontrolada, o marco do câncer.

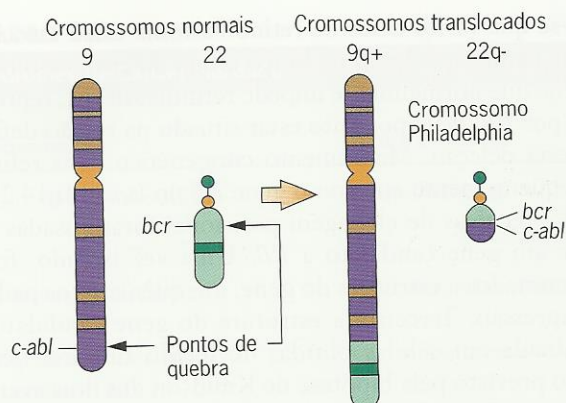
sas mutações deletérias ocorre em um oncogene celular. Assim, este grupo de genes desempenha um papel importante na etiologia do câncer humano.

### Rearranjos Cromossômicos e Câncer

Alguns tipos de câncer humano estão associados a rearranjos cromossômicos. Por exemplo, leucemia mieloide crônica (CML) está associada a uma anomalia do cromossomo 22. Este cromossomo anormal foi originalmente descoberto na cidade de Philadelphia e portanto denominado *cromossomo*

*Philadelphia*. Inicialmente, pensou-se que era uma simples deleção em seu braço longo. Entretanto, análise subsequente usando técnicas moleculares mostrou que o cromossomo Philadelphia de fato resulta de uma translocação recíproca entre os cromossomos 9 e 22. (Veja discussão geral sobre translocações no Cap. 6.) Na translocação Philadelphia, a ponta do braço longo do cromossomo 9 foi unida ao corpo do cromossomo 22, e a parte distal do braço longo do cromossomo 22 foi unida ao corpo do cromossomo 9 (Fig. 24.6). O ponto de quebra da translocação no cromossomo 9 é no oncogene *c-abl*, que codifica uma tirosina-cinase, e o





**Fig. 24.6** ■ A translocação recíproca envolvida no cromossomo Philadelphia associado a leucemia mieloide crônica.

ponto de quebra no cromossomo 22 é em um gene chamado *bcr*. Por meio da translocação, os genes *bcr* e *c-abl* foram fisicamente unidos, criando um gene de fusão cujo produto polipeptídico possui a ponta amino da proteína *bcr* e a ponta carboxi da proteína *abl*. Embora não seja compreendido exatamente o motivo, este polipeptídeo de fusão faz com que células sanguíneas se tornem cancerosas. O mecanismo pode envolver a atividade de tirosina-cinase da proteína *c-Abl*, que é rigidamente controlada em células normais mas é desregulada em células que produzem o polipeptídeo de fusão. De fato, a função da tirosina-cinase da proteína *c-Abl* foi ativada constitutivamente pelo gene de fusão *bcr/c-abl*. Essa fusão é portanto um ativador dominante da tirosina-cinase de *c-Abl*. A desregulação da tirosina-cinase de *c-Abl* resulta em fosforilação anormal de outras proteínas, incluindo algumas que estão envolvidas em controle do ciclo celular. Em seu estado fosforilado, tais proteínas fazem com que as células cresçam e se dividam descontroladamente.

O linfoma de Burkitt é outro exemplo de câncer de leucócitos associado a translocações recíprocas. Essas translocações invariavelmente envolvem o cromossomo 8 e um dentre três cromossomos (2, 14 e 22) que levam genes codificantes dos polipeptídeos que formam imunoglobulinas (também conhecidos como anticorpos; veja Cap. 23). Translocações

envolvendo os cromossomos 8 e 14 são as mais comuns (Fig. 24.7). Nessas translocações, o oncogene *c-myc* no cromossomo 8 é justaposto aos genes das cadeias pesadas de imunoglobulinas no cromossomo 14. Esse rearranjo resulta na hiperexpressão do oncogene *c-myc* em células que produzem cadeias pesadas de imunoglobulina, isto é, linfócitos B, e essa hiperexpressão faz com que tais células se tornem cancerosas.

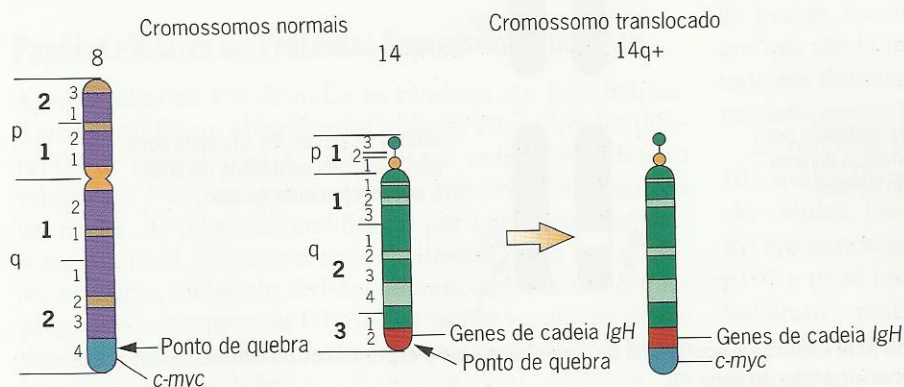
### PONTOS IMPORTANTES

- Alguns vírus levam genes (oncogenes) que podem induzir a formação de tumores em animais.
- Os oncogenes virais são homólogos aos genes celulares (protoncogenes), que podem induzir tumores quando são hiperexpressos ou quando estão mutados para produzir produtos protéicos anormalmente ativos.
- Mutações em protoncogenes promovem ativamente proliferação celular.
- Alguns cânceres estão associados a rearranjos cromossômicos que acentuam a expressão de protoncogenes ou que alteram a natureza de seus produtos protéicos.

## GENES SUPRESSORES TUMORAIS

Muitos cânceres envolvem a inativação de genes cujos produtos têm papéis importantes na regulação do ciclo celular.

Os alelos normais de genes tais como *c-ras* e *c-myc* produzem proteínas que regulam o ciclo celular. Quando tais genes são hiperexpressos ou quando produzem proteínas que funcionam como ativadores dominantes, a célula é predisposta a tornar-se cancerosa. Entretanto, o desenvolvimento total de um estado canceroso geralmente requer mutações adicionais, e tipicamente essas mutações afetam genes que normalmente estão envolvidos no controle do crescimento celular. Tais mutações definem, portanto, uma segunda classe de genes relacionados ao câncer, os antioncogenes, ou, como são mais geralmente chamados, genes supressores tumorais.



**Fig. 24.7** ■ Uma translocação recíproca envolvida no linfoma de Burkitt. Apenas o cromossomo translocado (14q+) que leva tanto o oncogene *c-myc* quanto os genes de cadeia pesada de imunoglobulina (*IgH*) é mostrado.



## Cânceres Herdados e a Hipótese de Knudson de Dois Eventos

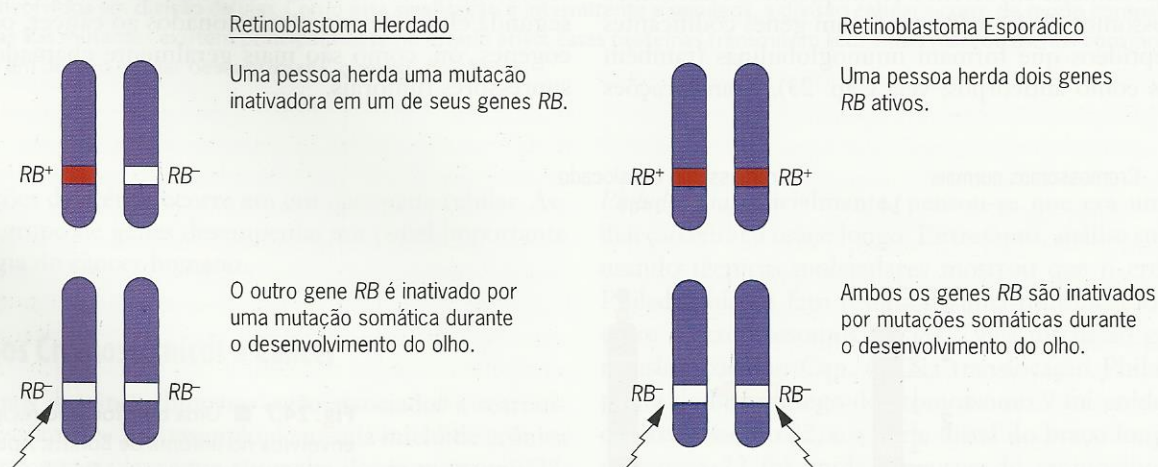
Muitos dos genes supressores tumorais foram inicialmente descobertos pela análise de cânceres raros em que uma predisposição para desenvolver o câncer segue um padrão de herança dominante. Essa predisposição se deve a heteroziguidade de uma mutação herdada de perda de função no gene supressor tumoral. O câncer só se desenvolve se ocorrer uma segunda mutação nas células somáticas e se essa mutação inativar a função do alelo tipo selvagem do gene supressor tumoral. Assim, o desenvolvimento do câncer requer duas mutações de perda de função, isto é, dois “eventos” de inativação, um em cada uma das duas cópias do gene supressor tumoral.

Em 1971, Alfred Knudson propôs esta explicação para a ocorrência de *retinoblastoma*, um raro câncer de olho infantil. Na maioria das populações humanas, a incidência de retinoblastoma é de cerca de 5 em 100.000 crianças. Análise de heredogramas indica que aproximadamente 40% dos casos envolvem uma mutação herdada que predispõe o indivíduo a desenvolver o câncer. Os outros 60% dos casos não podem ser correlacionados a uma mutação específica herdada. Esses casos não herdados são ditos *esporádicos*. Com base em análises estatísticas, Knudson propôs que tanto os casos herdados quanto os esporádicos de retinoblastoma ocorrem porque as duas cópias de um determinado gene foram inativadas (Fig. 24.8). Nos casos herdados, uma das mutações inativadoras foi transmitida através da linhagem germinativa e a outra ocorre durante o desenvolvimento dos tecidos somáticos do olho. Nos casos esporádicos, ambas as mutações inativadoras ocorrem durante o desenvolvimento do olho. Assim, em ambos os tipos de retinoblastoma, são necessários dois “eventos” mutacionais para inativar um gene que normalmente funciona para suprimir a formação de tumor no olho.

Achados subsequentes de pesquisa verificaram a exatidão da hipótese de dois eventos de Knudson. Primeiro, desco-

briu-se que vários casos de retinoblastoma estão associados a uma pequena deleção no braço longo do cromossomo 13. O gene que normalmente impede retinoblastoma, representado por *RB*, deve portanto estar situado na região definida por esta deleção. Mapeamento citogenético mais refinado subsequentemente colocou o gene *RB* no *locus* 13q14.2. Segundo, técnicas de clonagem posicional foram usadas para isolar um gene candidato a *RB*. Uma vez isolado, foram determinados a estrutura do gene, a sequência e os padrões de expressão. Terceiro, a estrutura do gene candidato foi examinada em células obtidas de tecido tumoral ocular. Como previsto pela hipótese de Knudson dos dois eventos, ambas as cópias deste gene estavam inativadas em células do retinoblastoma. Assim, o gene candidato parece ser o autêntico gene *RB*. Finalmente, experimentos de cultura celular demonstraram que um cDNA do alelo tipo selvagem do gene candidato podia reverter as propriedades cancerosas de células tumorais cultivadas. Tais experimentos de reversão de câncer provaram sem nenhuma dúvida que o gene candidato era o autêntico gene supressor tumoral *RB*. Foi verificado que o produto protéico deste gene, indicado por pRB, é uma proteína ubiquamente expressa que interage com uma família de fatores de transcrição envolvidos na regulação do ciclo celular.

A hipótese de Knudson dos dois eventos foi então aplicada a outros cânceres herdados, incluindo tumor de Wilms, síndrome Li-Fraumeni, neurofibromatose, doença de von Hippel-Lindau e alguns tipos de câncer de cólon e de mama (Quadro 24.2). Em cada caso, está envolvido um gene supressor tumoral diferente. Por exemplo, no tumor de Wilms, um câncer do sistema urogenital, o gene supressor tumoral relevante é o gene *WT1* situado no braço curto do cromossomo 11. Na neurofibromatose, uma doença caracterizada por tumores benignos e lesões de pele, é o gene *NF1* situado no braço longo do cromossomo 17; e na polipose adenomatosa familiar, uma condição caracterizada pela ocorrência de vários tumores no cólon, é o gene



**Fig. 24.8** ■ Hipótese de dois eventos de Knudson para explicar a ocorrência de casos herdados e esporádicos de retinoblastoma. São necessárias duas mutações inativadoras para eliminar a função do gene *RB*.



## QUADRO 24.2

## Síndromes de Câncer Herdado

| Síndrome                                           | Tumor Primário                  | Gene                                                     | Local Cromossômico           | Função Protéica Proposta                          |
|----------------------------------------------------|---------------------------------|----------------------------------------------------------|------------------------------|---------------------------------------------------|
| Retinoblastoma familiar                            | Retinoblastoma                  | <i>RB</i>                                                | 13q14.3                      | Ciclo celular e regulação transcricional          |
| Síndrome Li-Fraumeni                               | Sarcomas, câncer de mama        | <i>TP53</i>                                              | 17p13.1                      | Fator de transcrição                              |
| Polipose adenomatosa familiar (FAP)                | Câncer colorretal               | <i>APC</i>                                               | 5q21                         | Regulação de $\beta$ -catenina                    |
| Câncer colorretal não-polipose hereditário (HNPCC) | Câncer colorretal               | <i>MSH2</i><br><i>MLH1</i><br><i>PMS1</i><br><i>PMS2</i> | 2p16<br>3p21<br>2q32<br>7p22 | Reparo de pareamento errado no DNA                |
| Neurofibromatose tipo 1                            | Neurofibromas                   | <i>NF1</i>                                               | 17q11.2                      | Regulação de sinalização mediada por Ras          |
| Neurofibromatose tipo 2                            | Neuromas acústicos, meningiomas | <i>NF2</i>                                               | 22q12.2                      | Ligação de proteínas da membrana ao citoesqueleto |
| Tumor de Wilms                                     | Tumor de Wilms                  | <i>WT1</i>                                               | 11p13                        | Repressor transcricional                          |
| Câncer de mama familiar 1                          | Câncer de mama                  | <i>BRCA1</i>                                             | 17q21                        | Reparo do DNA                                     |
| Câncer de mama familiar 2                          | Câncer de mama                  | <i>BRCA2</i>                                             | 13q12                        | Reparo do DNA                                     |
| Doença de von Hippel-Lindau                        | Câncer renal                    | <i>VHL</i>                                               | 3p25                         | Regulação de alongamento transcricional           |
| Melanoma familiar                                  | Melanoma                        | <i>p16</i>                                               | 9p21                         | Inibidor de CDK                                   |
| Ataxia-telangiectasia                              | Linfoma                         | <i>ATM</i>                                               | 11q22                        | Reparo do DNA                                     |
| Síndrome Bloom                                     | Tumores sólidos                 | <i>BLM</i>                                               | 15q26.1                      | DNA-helicase                                      |

Fonte: Fearon, E. R. 1997. Human cancer syndromes: clues to the origin and nature of cancer. *Science* 278:1043-1050.

*APC* situado no braço longo do cromossomo 5. Como o retinoblastoma, estas três doenças são raras, e apenas uma fração dos casos observados envolve uma mutação herdada no gene supressor tumoral relevante. Os outros casos são causados ou por duas mutações somáticas independentes neste gene ou por mutações em outros genes supressores tumorais ainda não identificados.

### Papéis Celulares de Proteínas Supressoras Tumorais

Apenas cerca de 1% de todos os cânceres são hereditários. Entretanto, foram identificadas mais de 20 síndromes diferentes de câncer herdado, e em quase todas elas o defeito subjacente é em um gene supressor tumoral e não em um oncogene. As proteínas codificadas por esses genes supressores tumorais funcionam em uma diversa gama de processos celulares, incluindo divisão, diferenciação, morte celular programada e reparo do DNA. Nas seções seguintes, discutimos algumas das proteínas supressoras tumorais que foram intensamente estudadas.

### pRB

Pesquisas recentes revelaram que a proteína supressora tumoral RB tem um papel central na regulação do ciclo celular. Embora o gene *RB* tenha sido descoberto por sua associação com retinoblastoma, mutações neste gene também estão associadas a outros tipos de câncer, incluindo carcinomas de pequenas células do pulmão, osteossarcomas e carcinomas de bexiga, cervical e da próstata. Além disso, camundongos que são homozigotos para uma mutação nocaute de *RB* morrem durante o desenvolvimento embrionário. Assim, o produto gênico RB é essencial para a vida.

O produto do gene RB é uma proteína nuclear com 105 quilodáltons que está envolvida na regulação do ciclo celular. Foram encontrados dois genes homólogos a *RB* em genomas de mamíferos, e seus produtos protéicos, p107 e p130 (cada um denominado por sua massa em quilodáltons), podem também ter papéis principais na regulação do ciclo celular. Não se sabe se tumores humanos têm mutações inativadoras em nenhum destes dois genes,

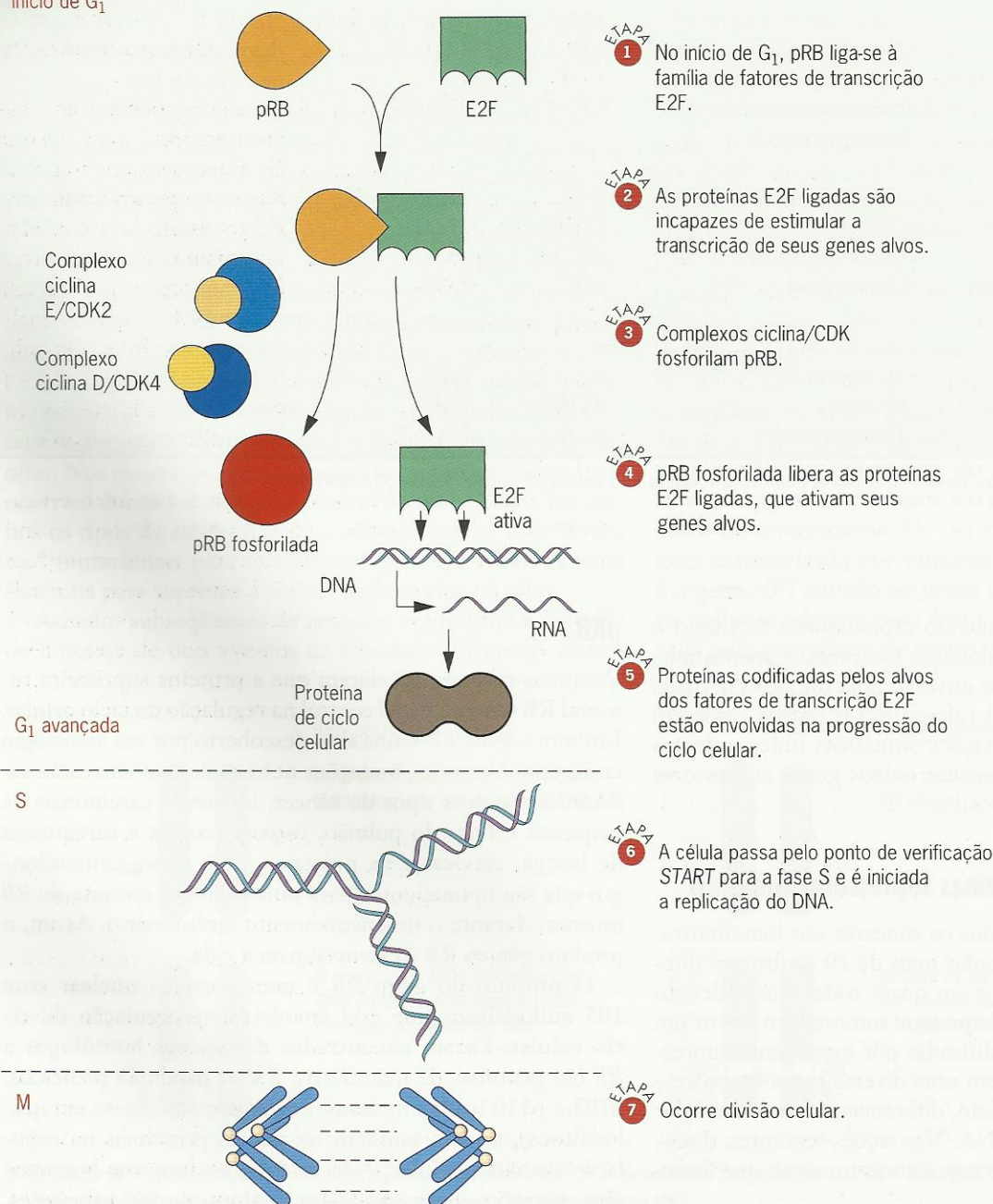


e camundongos homozigotos para uma mutação nocaute em qualquer um deles não apresentam fenótipos anormais. Entretanto, camundongos que são homozigotos para mutações nocaute em ambos estes genes morrem logo após o nascimento. Assim, os membros p107 e p130 da família RB de proteínas juntos estão envolvidos em processos celulares importantes.

Análises moleculares e bioquímicas elucidaram o papel de pRB na regulação do ciclo celular (Fig. 24.9). No início da fase  $G_1$  do ciclo celular, pRB liga-se às proteínas E2F, uma família de fatores de transcrição que controla a expressão de vários genes cujos produtos movem a célula ao longo de seu ciclo. Quando fatores de transcrição E2F estão ligados

a pRB, eles não podem ligar-se a seqüências acentuadoras específicas em seus genes alvos. Conseqüentemente, os fatores do ciclo celular codificados por estes genes não são produzidos, e a maquinaria para síntese de DNA e divisão celular permanece quiescente. Mais adiante em  $G_1$ , pRB é fosforilada pela ação de cinases dependentes de ciclina. Neste estado alterado, pRB libera os fatores de transcrição E2F ligados a ela. Esses fatores de transcrição liberados estão então livres para ativar seus genes alvos, que codificam proteínas com a função de induzir a célula a progredir para a fase S e para a mitose. Após a mitose, pRB é desfosforilada, e cada uma das células filhas entra na fase quiescente de um novo ciclo celular.

Início de  $G_1$



**Fig. 24.9** ■ Papel de pRB na progressão do ciclo celular. Por meio de sua interação negativa com os fatores de transcrição E2F, pRB pára o ciclo celular na fase  $G_1$ . A fosforilação de pRB pelos complexos ciclina/CDK liberta proteínas E2F para ativar seus genes alvos, que codificam proteínas que são instrumentais em mover a célula além do ponto de verificação START para a fase S.



Esta progressão ordenada e rítmica pelo ciclo celular é perturbada em células cancerosas. Em muitos tipos de câncer – não só no retinoblastoma –, ambas as cópias do gene *RB* foram inativadas, seja por deleções ou por mutações que prejudicam ou abolem a capacidade da proteína RB de se ligar a fatores de transcrição E2F. A incapacidade de pRB de se ligar a esses fatores de transcrição os deixa livres para ativar seus genes alvos, movimentando assim a maquinaria de síntese de DNA e divisão celular. De fato, um dos freios naturais do processo de divisão celular foi liberado. Na ausência desse freio, as células têm uma tendência a mover-se rapidamente por seu ciclo. Se outros freios do ciclo celular falham, as células dividem-se incessantemente para formar tumores.

### p53

A proteína supressora tumoral p53, com 53 quilodáltons, foi descoberta por seu papel na indução de cânceres por alguns vírus com DNA. Esta proteína é codificada por um gene supressor tumoral chamado *TP53*. Mutações herdadas em *TP53* estão associadas à síndrome Li-Fraumeni, uma condição dominante rara na qual pode desenvolver-se um dentre vários tipos de câncer. Mutações somáticas que inativam ambas as cópias do gene *TP53* também estão associadas a uma variedade de cânceres. De fato, tais mutações são encontradas na maioria de todos os tumores humanos. Perda de função de p53 é portanto uma etapa chave na carcinogênese.

A proteína p53 é um fator de transcrição com 393 aminoácidos que consiste em três domínios diferentes: um domínio N-terminal de ativação de transcrição (TAD), um domínio cerne de ligação central ao DNA (DBD) e um domínio C-terminal de homo-oligomerização (OD) (Fig. 24.10a). A maioria das mutações que inativam p53 estão localizadas no DBD. Tais mutações evidentemente prejudicam ou abolem a capacidade de p53 de se ligar a seqüências específicas do DNA que estão inseridas em seus genes alvos, impedindo assim a ativação transcricional desses genes. Assim, mutações que alteram DBD são tipicamente mutações recessivas de perda de função. Outros tipos de mutações são encontradas na parte OD do polipeptídeo. Moléculas de p53 com estes tipos de mutações dimerizam com polipeptídeos p53 tipo selvagem e impedem que os polipeptídeos tipo selvagem funcionem como ativadores transcricionais. Assim, mutações que alteram OD têm um efeito *dominante negativo* sobre a função de p53.

A proteína p53 tem um papel central nas respostas celulares a estresse (Fig. 24.10b). Em células normais, o nível de p53 é baixo, mas, quando as células são tratadas com um agente que danifica o DNA, tal como radiação, o nível de p53 aumenta acentuadamente. Esta resposta a dano no DNA é mediada por uma via que diminui a degradação de p53. Em resposta a dano no DNA, p53 é fosforilada, convertendo-se em uma forma estável e ativa. Uma vez ativada, p53 estimula a transcrição de genes cujos produtos detêm o

ciclo celular, permitindo assim que o DNA danificado seja reparado, ou ativa outro conjunto de genes cujos produtos no final causam a morte da célula danificada.

Um fator proeminente na resposta que interrompe o ciclo celular é p21, uma proteína codificada por um gene que é ativado pelo fator de transcrição p53. A proteína p21 é um inibidor dos complexos protéicos ciclina/CDK. Quando p21 é sintetizada em resposta a estresse celular, os complexos ciclina/CDK são inativados e o ciclo celular é detido. Durante essa interrupção, o DNA danificado da célula pode ser reparado. Assim, p53 é responsável por ativar um freio no ciclo celular, e esse freio permite que a célula mantenha sua integridade genética. Células que não possuem p53 funcional têm dificuldade em aplicar esse freio. Se tais células progredirem para as divisões subseqüentes, mutações adicionais que as aumentam podem acumular-se. Inativação mutacional de p53 é, portanto, geralmente uma etapa importante na via para o câncer.

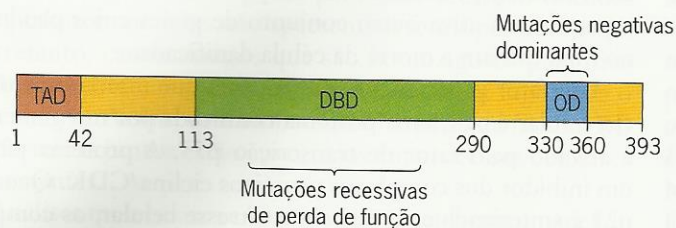
A proteína p53 também pode mediar outra resposta a estresse celular. Em vez de orquestrar esforços para reparar danos dentro de uma célula, a p53 pode disparar uma resposta suicida em que a célula danificada é programada para destruição. O modo pelo qual a p53 programa morte celular não é bem compreendido. Um mecanismo parece envolver o produto protéico do gene *BAX*. A proteína BAX é um antagonista de outra proteína chamada BCL-2, que normalmente suprime a via apoptótica, ou morte celular. Quando o gene *BAX* é ativado por p53, seu produto protéico libera a proteína BCL-2 de seu modo supressor. Essa liberação então abre a via apoptótica, e a célula continua em sua própria destruição.

Curiosamente, a proteína p53 não parece exercer um papel significativo na morte celular programada que ocorre durante embriogênese. Camundongos que são homozigotos para mutações nocaute no *TP53* desenvolvem-se normalmente, embora tenham tendência a desenvolver tumores à medida que envelhecem. Assim, a despeito de seu papel central em regular respostas celulares a estresse, a p53 não parece influenciar o curso do desenvolvimento embrionário.

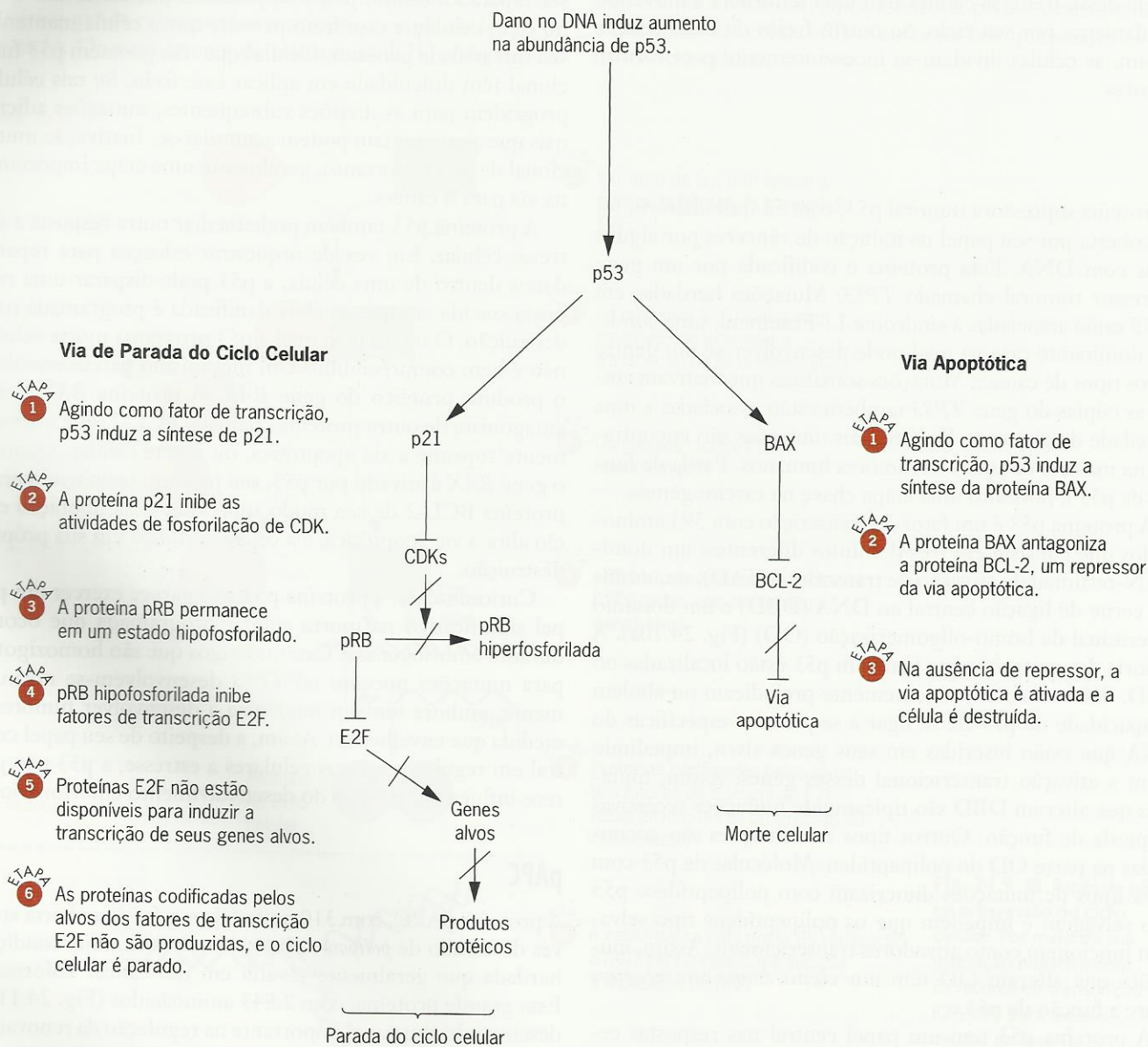
### pAPC

A proteína pAPC, com 310 quilodáltons, foi descoberta através do estudo de *polipose adenomatosa do cólon*, uma condição herdada que geralmente resulta em um câncer colorretal. Esta grande proteína, com 2.843 aminoácidos (Fig. 24.11a), desempenha um papel importante na regulação da renovação de células no revestimento, ou epitélio, do intestino grosso. Embora os mecanismos que regulam este processo não sejam totalmente compreendidos, informações atuais sugerem que pAPC controla a proliferação e a diferenciação de células no epitélio do intestino. Quando a função de pAPC é perdida, as células que geram as projeções digitiformes do epitélio intestinal permanecem em um estado indiferenciado. A medida que tais células continuam a se dividir, elas produzem





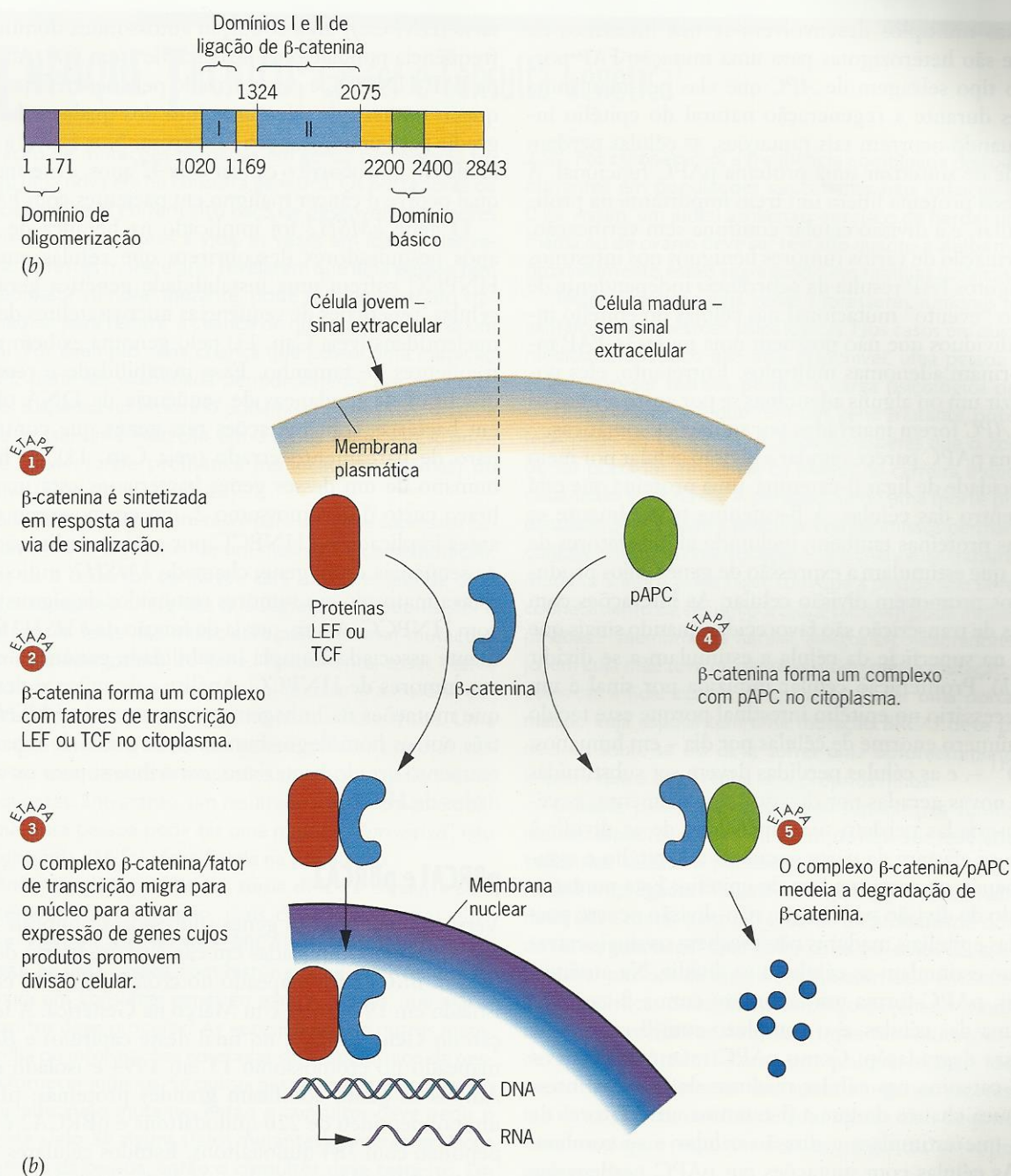
(a)



(b)

**Fig. 24.10** ■ (a) Domínios principais dentro de p53. TAD = domínio de ativação de transcrição; DBD = domínio de ligação ao DNA; OD = domínio de oligomerização. Os números referem-se a posições de aminoácidos no polipeptídeo. (b) Papel de p53 na resposta celular a dano no DNA. Foram identificadas duas vias de respostas. Dentro de cada via, a ponta da seta indica uma influência positiva ou uma mudança direcional (p. ex., uma proteína é sintetizada ou fosforilada, uma proteína catalisa uma reação ou um gene é expresso) e uma seta romba indica uma influência negativa (p. ex., repressão de síntese de proteínas ou atividade da proteína ou repressão de uma via). Uma barra por cima de uma seta indica que a influência, positiva ou negativa, está bloqueada.





**Fig. 24.11** ■ (a) Domínios principais dentro de pAPC. Os números referem-se a posições de aminoácidos no polipeptídeo. (b) Papel de pAPC no controle do ciclo celular. A proteína pAPC influencia a progressão pelo ciclo celular interagindo com  $\beta$ -catenina, uma proteína que pode ativar os fatores de transcrição LEF ou TCF. Em células jovens (esquerda), um sinal extracelular ativa esses fatores de transcrição e a divisão celular é estimulada. Em células maduras (direita), interações de pAPC e  $\beta$ -catenina impedem que os fatores de transcrição sejam ativados, e a divisão celular é inibida.

mais de seu tipo, e o aumento resultante do número de células causa a formação de muitos pequenos tumores benignos no epitélio intestinal. Esses tumores são chamados *pólipos* ou *adenomas*, e a predisposição a formá-los é herdada como uma condição autossômica dominante rara chamada *polipose adenomatosa familiar (FAP)*. Nos países ocidentais, sua frequência populacional é de cerca de 1 em 7.000.

Pacientes com FAP desenvolvem múltiplos adenomas durante a adolescência e o início da segunda década de vida. Embora os adenomas sejam inicialmente benignos, há uma grande probabilidade de que pelo menos um deles se torne um tumor maligno. Assim, em uma idade relativamente jovem – nos EUA, a média é de 42 –, portadores de uma mutação FAP desenvolvem um câncer colorretal manifesto.



Adenomas múltiplos desenvolvem-se nos intestinos de pessoas que são heterozigotas para uma mutação FAP porque o alelo tipo selvagem de *APC* que elas possuem muta várias vezes durante a regeneração natural do epitélio intestinal. Quando ocorrem tais mutações, as células perdem a capacidade de sintetizar uma proteína pAPC funcional. A ausência dessa proteína libera um freio importante na proliferação celular, e a divisão celular continua sem verificação. Assim, a formação de vários tumores benignos nos intestinos de heterozigotos FAP resulta da ocorrência independente de um segundo "evento" mutacional nas células do epitélio intestinal. Indivíduos que não possuem uma mutação FAP raramente formam adenomas múltiplos. Entretanto, eles podem produzir um ou alguns adenomas se por acaso ambos os seus genes *APC* forem inativados por mutações somáticas.

A proteína pAPC parece regular a divisão celular por meio de sua capacidade de ligar  $\beta$ -catenina, uma proteína que está presente dentro das células. A  $\beta$ -catenina naturalmente se liga a outras proteínas também, incluindo alguns fatores de transcrição que estimulam a expressão de genes cujos produtos protéicos promovem divisão celular. As interações com estes fatores de transcrição são favorecidas quando sinais que interferem na superfície da célula a estimulam a se dividir (Fig. 24.11b). Proliferação celular induzida por sinal é um processo necessário no epitélio intestinal porque este tecido perde um número enorme de células por dia – em humanos, cerca de  $10^{11}$  –, e as células perdidas devem ser substituídas por células novas geradas por divisão. Normalmente, as células recém-criadas perdem sua capacidade de se dividir à medida que se afastam da parte geradora de epitélio e assumem seus papéis na parte madura do epitélio. Esta mudança de um estado de divisão para um de não-divisão ocorre porque as células epiteliais maduras não recebem os sinais extracelulares que estimulam as células a se dividir. Na ausência desses sinais, pAPC forma um complexo com a  $\beta$ -catenina no citoplasma das células, e o complexo com  $\beta$ -catenina é marcado para degradação. Como pAPC mantém baixos os níveis de  $\beta$ -catenina nas células maduras do epitélio intestinal, há pouca chance de que a  $\beta$ -catenina ative fatores de transcrição que estimulam a divisão celular e se combine com eles. As células com mutações em pAPC perdem sua capacidade de controlar os níveis de  $\beta$ -catenina. Sem esse controle, elas conservam seu vigor para divisão e não se diferenciam apropriadamente em células epiteliais maduras. O resultado é que um tumor benigno começa a formar-se no revestimento intestinal. Assim, moléculas normais de pAPC têm um papel importante em suprimir a formação de tumores no intestino.

## phMSH2

A proteína phMSH2 é o homólogo humano de uma proteína de reparo do DNA chamada MutS encontrada em bactérias e leveduras. Seu envolvimento em câncer humano foi elucidado pelo estudo do *câncer colorretal não-polipose hereditário* (HNPCC), uma condição autossômica dominante com

freqüência populacional de cerca de 1 em 500. Ao contrário da FAP, HNPCC é caracterizado pela ocorrência de um pequeno número de adenomas, um dos quais acaba por progredir para uma condição cancerosa. Nos EUA, a média de idade em que ocorre o câncer é de 42 anos, a mesma idade na qual ocorre o câncer maligno em pacientes com FAP.

O gene *bMSH2* foi implicado na herança de HNPCC após pesquisadores descobrirem que células em tumores HNPCC sofrem uma instabilidade genética geral. Nessas células, repetições de seqüências microssatélites de di- e trinucleotídeos (veja Cap. 13) pelo genoma exibem mudanças freqüentes de tamanho. Essa instabilidade é remanescente dos tipos de mudanças de seqüência de DNA observadas em bactérias com mutações nos genes que controlam reparo de pareamento errado (veja Cap. 13). O homólogo humano de um desses genes bacterianos está mapeado no braço curto do cromossomo 2, um cromossomo que já foi antes implicado no HNPCC por análise de ligação. Análise de seqüência deste gene, chamado *bMSH2*, indicou que ele estava inativado em tumores removidos de alguns pacientes com HNPCC. Assim, perda de função de *bMSH2* foi causalmente associada à ampla instabilidade genômica observada em tumores de HNPCC. Análise subsequente demonstrou que mutações na linhagem germinativa do *bMSH2*, ou em três outros homólogos humanos de genes de reparo de pareamento errado bacteriano, contribuem para os casos herdados de HNPCC.

## pBRCA1 e pBRCA2

Versões mutantes dos genes supressores tumorais *BRCA1* e *BRCA2* foram implicadas em câncer hereditário de mama e ovário. *BRCA1* foi mapeado no cromossomo 17 em 1990 e isolado em 1994 (veja Um Marco na Genética: A Identificação do Gene *BRCA1*, no final deste capítulo) e *BRCA2* foi mapeado no cromossomo 13 em 1994 e isolado em 1995. Ambos os genes codificam grandes proteínas; pBRCA1 é um polipeptídeo de 220 quilodáltons e pBRCA2 é um polipeptídeo com 384 quilodáltons. Estudos celulares e bioquímicos mostraram que cada uma destas proteínas está situada dentro do núcleo de células normais, e cada uma contém um domínio hipotético de ativação transcricional. Entretanto, não sabemos se tais proteínas são de fato fatores de transcrição. As proteínas pBRCA1 e pBRCA2 também contêm um domínio que lhes permite interagir fisicamente com outras proteínas, em particular com pRAD51, um homólogo eucariótico da proteína bacteriana de reparo do DNA conhecida como RecA. Assim, pBRCA1 e pBRCA2 podem estar envolvidas em um dos muitos sistemas de reparo do DNA danificado em células humanas.

Tanto pBRCA1 quanto pBRCA2 desempenham funções importantes dentro das células. Camundongos que são homozigotos para uma mutação nocaute em um destes genes morrem cedo durante a embriogênese. Ainda não está claro





## ENFOQUE TÉCNICO: Câncer e Consulta Genética<sup>1</sup>

A identificação de mutações herdadas em genes supressores tumorais abriu uma nova era na consulta genética. Os portadores de tais mutações em geral correm alto risco de desenvolver tumores potencialmente ameaçadores à vida, às vezes em idade relativamente nova. Se testes moleculares revelarem que uma pessoa tem um gene supressor tumoral mutante, pode ser administrado tratamento médico para reduzir a chance de que ela desenvolva um câncer letal. Por exemplo, uma criança que possui uma mutação no gene *APC* pode ser examinada periodicamente por endoscopia, e lesões suspeitas no intestino podem ser removidas, ou uma mulher que possui uma mutação em um dos genes *BRCA* pode sofrer uma mastectomia profilática (remoção das mamas) ou ooforectomia (remoção dos ovários).

Um resultado negativo de um teste para um gene supressor tumoral mutante seria, logicamente, um motivo de comemoração, pelo menos se o teste for confiável. Para um gene grande com muitos alelos mutantes diferentes que segregam na população, é difícil imaginar um teste não muito caro para detectar mutações situadas em qualquer ponto do gene. Tipicamente, esses testes são baseados na reação em cadeia da polimerase, e a maioria deles é projetada para detectar alelos mutantes específicos. Uma pessoa que está em risco de ter um gene supressor tumoral mutante pode ser testada para as mutações conhecidas, pelo menos as mais frequentes. Entretanto, um resultado negativo não é definitivo porque essa pessoa pode ter uma mutação "privativa", isto é, aquela que ainda não foi identificada na população.

A existência de alelos privativos torna difícil a consulta para cânceres herdados. Por exemplo, mais de 300 mutações diferentes foram identificadas no gene *BRCA1*, e cerca de 50% delas são privativas. Se uma pessoa com história familiar de câncer de mama procura um consultor genético para avaliação, que mutações o consultor deve procurar? Às vezes, dados de outros membros da família ou informações coletadas do grupo étnico da pessoa podem fornecer indícios. Se outras pessoas na família tiverem um determinado alelo mutante, então o consultor deve pedir o teste para este alelo. Se alguns alelos mutantes são característicos do grupo étnico da pessoa, então o consultor deve testá-los. Em populações de judeus ashkenazi, por exemplo, alguns alelos mutantes *BRCA1* e *BRCA2* têm frequências alélicas tão altas quanto

2,5%. Por comparação, a frequência combinada de todos os alelos mutantes em populações caucasianas não judaicas é de apenas 0,1%. Assim, um judeu ashkenazi em risco de herdar um câncer de mama ou de ovário deve ser testado quanto a alelos mutantes que provavelmente estão segregando em famílias judaicas ashkenazi.

Testes genéticos para genes supressores tumorais mutantes levantam várias questões psicológicas. Nos casos em que tratamento médico terapêutico não está disponível, uma pessoa poderia escolher não ser testada porque a carga psicológica de viver com o conhecimento de que tem um gene mutante potencialmente letal pode ser insuportável. Pode-se esperar que o conhecimento de ser portador influencie os planos de carreira e decisões sobre casamento e ter filhos. A perspectiva de uma morte precoce pode dissuadir uma pessoa a procurar compromissos permanentes – um cônjuge, filhos ou uma vocação –, e a chance de transmitir um alelo mutante a filhos pode fazer com que a pessoa não se reproduza. O conhecimento de ser um portador também pode influenciar outras pessoas – membros da família, amigos e colaboradores. Uma filha jovem cuja mãe teve um teste positivo para a mutação *BRCA1* deve começar a lidar com a perspectiva de ser uma portadora, e um marido cuja esposa tem uma mutação *BRCA1* deve participar da decisão sobre se ela deve sofrer uma ooforectomia profilática e impedir que o casal tenha seus próprios filhos.

O teste para genes mutantes supressores tumorais levanta muitos aspectos éticos. A quem devem ser revelados os resultados dos testes? o paciente? a família do paciente? genitores? filhos? empregador? locatário? seguradora? Que medidas a sociedade deve tomar para salvaguarda da privacidade dos resultados de testes genéticos? Que políticas os governos devem adotar para proteger pessoas de discriminação com base em seus genótipos? Como políticas de seguro e emprego devem ser modificadas? Os direitos reprodutivos de pessoas que têm mutações prejudiciais devem ser limitados? Como em qualquer avanço tecnológico, a capacidade de detectar mutações em genes supressores tumorais deixa-nos com muitas dúvidas sobre como devemos proceder. Atualmente, as respostas a estas perguntas estão longe de ser esclarecidas.

<sup>1</sup>Ponder, Bruce. 1997. Genetic testing for cancer risk. *Science* 278: 1050-1054.

que papéis as proteínas mutantes pBRCA1 e pBRCA2 desempenham na carcinogênese humana. Possivelmente elas comprometem a capacidade da célula de detectar ou reparar DNA danificado.

Mutações nos genes *BRCA1* e *BRCA2* respondem por cerca de 7% de todos os casos de câncer de mama e cerca de 10% de todos os casos de câncer de ovário nos EUA. Para cada gene, a predisposição a desenvolver esses cânceres é herdada como um alelo dominante com alta pene-

trância. Portadores têm um risco de 10 a 25 vezes maior que não portadores de desenvolver câncer de mama ou de ovário, e em algumas famílias o risco de desenvolver câncer de cólon ou próstata também está aumentado. Como muitas mutações inativadoras diferentes em *BRCA1* e *BRCA2* são encontradas na população humana, consulta genética para famílias que estão segregando essas mutações pode ser difícil (veja Enfoque Técnico: Câncer e Consulta Genética).



### PONTOS IMPORTANTES

- Os genes supressores tumorais foram descobertos por sua associação com cânceres herdados raros tais como retinoblastoma.
- Inativação mutacional de vários genes supressores tumorais é característica da maioria das formas de câncer.
- São necessários dois eventos mutacionais para eliminar ambas as cópias funcionais de um gene supressor tumoral dentro de uma célula.
- As proteínas codificadas por genes supressores tumorais têm papéis centrais na regulação do ciclo celular.

## VIAS GENÉTICAS PARA O CâNCER

Cânceres desenvolvem-se por acúmulo de mutações somáticas em protooncogenes e genes supressores tumorais.

Na maioria dos casos, a formação de um tumor maligno não é atribuível à ativação descontrolada de um único protooncogene ou à inativação de um único gene supressor tumoral. Em vez disso, a formação de tumor, o crescimento e a metástase geralmente dependem do acúmulo de mutações em vários genes diferentes. Assim, as vias genéticas para o câncer são diversas e complexas.

Podemos observar essa diversidade e complexidade na formação e no desenvolvimento de tipos diferentes de tumores. Por exemplo, tumores benignos do intestino grosso desenvolvem-se em pessoas com mutações inativadoras no gene *APC*. Entretanto, a progressão desses tumores para cânceres potencialmente letais requer mutações em vários outros genes. Esta via mutacional é resumida na Fig. 24.12a. Mutações inativadoras no gene *APC* iniciam o processo de formação de tumores causando o desenvolvimento de tecidos anormais dentro do epitélio intestinal. Esses tecidos anormais contêm células displásicas – células com formas incomuns e núcleos aumentados – que podem crescer em estágios iniciais de adenomas. Se o protooncogene *K-ras* é ativado em um desses adenomas, o adenoma pode crescer e desenvolver-se mais completamente. Mutações inativadoras em qualquer um dos vários genes supressores tumorais situados no braço longo do cromossomo 18 podem então induzir o adenoma a progredir mais e mutações inativadoras no gene supressor tumoral *TP53* no cromossomo 17 podem transformá-lo em um carcinoma de crescimento vigoroso. Mutações adicionais em genes supressores tumorais podem permitir que células do carcinoma invadam outros tecidos. Assim, não menos que sete mutações independentes (dois eventos inativadores no gene *APC*, uma mutação ativadora no gene *K-ras*, dois eventos inativadores em um gene supressor tumoral no cromossomo 18 e dois eventos inativadores no gene *TP53*) são necessárias para o desenvolvimento de um carcinoma intestinal, e ainda mais mutações provavel-

mente são necessárias para a metástase do carcinoma em direção a outras partes do corpo.

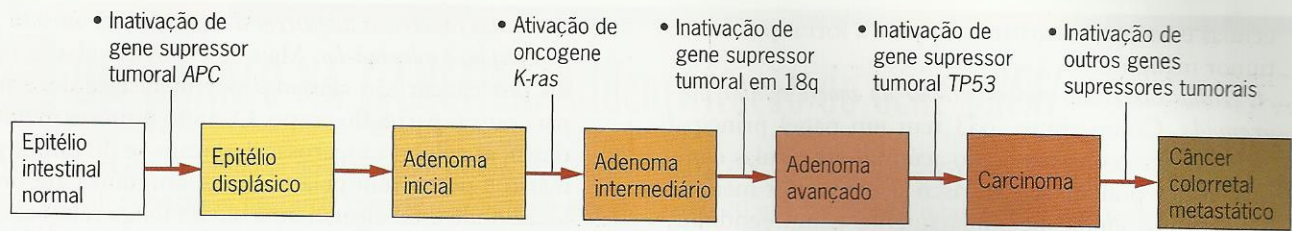
As vias genéticas para câncer de próstata e cérebro também foram elucidadas (Fig. 24.12b e Fig. 24.12c). Mutações em *HPC1*, um gene para câncer hereditário de próstata situado no braço longo do cromossomo 1, foram implicadas na origem de tumores de próstata. Mutações em outros genes supressores tumorais situados nos cromossomos 13, 16, 17 e 18 podem transformar tumores de próstata em cânceres metastáticos, e hiperexpressão do protooncogene *BCL-2* pode tornar esses cânceres imunes a terapia de privação de andrógenos, uma técnica padrão para o tratamento do câncer de próstata. O hormônio esteróide androgênico é necessário para a proliferação de células no epitélio prostático. Na ausência de andrógeno, tais células são programadas para morrer. Entretanto, células tumorais prostáticas podem adquirir a capacidade de sobreviver na ausência de andrógeno, provavelmente porque um excesso do produto gênico de *BCL-2* reprime a via de morte celular programada. Cânceres de próstata que progrediram para o estágio de independência de andrógeno são quase sempre fatais.

Glioblastomas são tumores que se desenvolvem de células gliais no cérebro. Glioblastomas primários geralmente ocorrem em pessoas com mais idade; eles crescem rapidamente e são quase sempre fatais. Glioblastomas secundários geralmente ocorrem em crianças e adultos jovens; eles se desenvolvem a partir de tumores preexistentes de baixo grau chamados astrocitomas e crescem lentamente, uma característica que os torna mais tratáveis do que glioblastomas primários. Dois protooncogenes foram implicados na formação de glioblastomas. O gene para o receptor do fator de crescimento epidérmico (EGFR) frequentemente é hiperexpresso em glioblastomas primários e o gene para a proteína fator de crescimento derivado de plaquetas (PDGF) frequentemente é hiperexpresso em glioblastomas secundários. Mutações em uma variedade de genes supressores tumorais, incluindo *TP53*, *RB* e *NF2*, foram implicadas na formação de glioblastomas, principalmente no desenvolvimento de glioblastomas secundários a partir de astrocitomas de baixo grau. Muitos outros genes supressores tumorais parecem estar envolvidos na formação de glioblastomas, mas eles ainda não foram identificados.

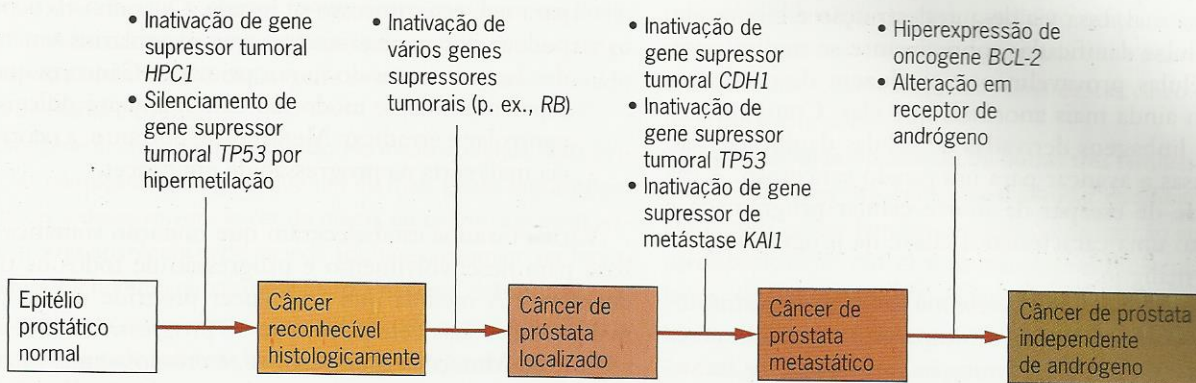
Douglas Hanahan e Robert Weinberg propuseram seis marcos das vias que levam ao câncer maligno:

1. As células cancerosas adquirem auto-suficiência na sinalização de processos que estimulam divisão e crescimento. Essa auto-suficiência pode surgir por mudanças nos fatores extracelulares que estimulam células a se dividir ou por mudanças em qualquer parte do sistema que transduz estes estímulos ou traduzem suas instruções em ação dentro da célula. No caso mais extremo, auto-suficiência ocorre quando células respondem a fatores de crescimento que elas mesmas produzem, criando assim um *feedback* positivo que estimula divisão celular incessante.

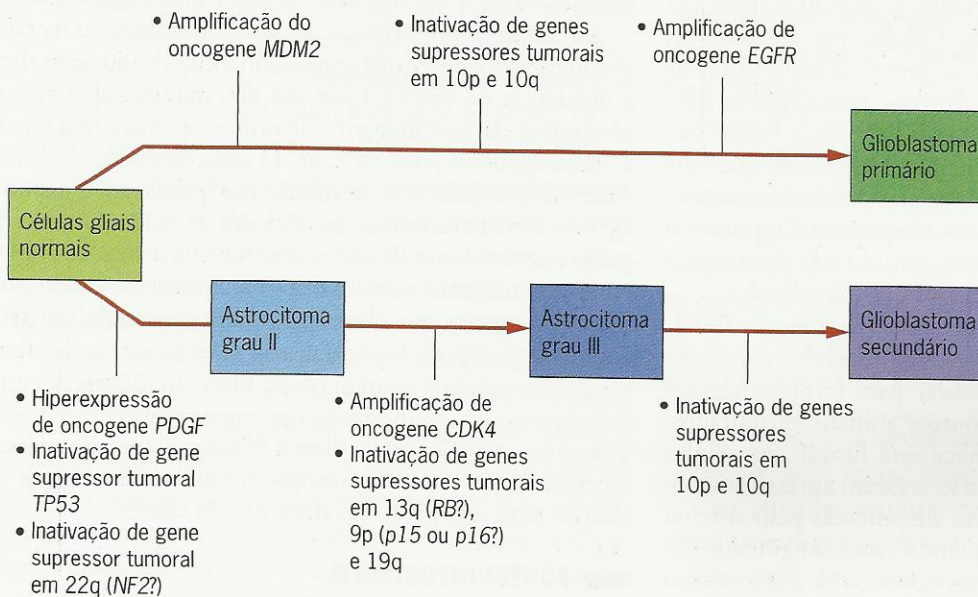




(a) Via de câncer colorretal metastático.



(b) Via de câncer de próstata independente de andrógeno.



(c) Vias para glioblastomas primário e secundário.

**Fig. 24.12** ■ Vias genéticas para o câncer.

2. As células cancerosas são anormalmente insensíveis a sinais que inibem o crescimento. A divisão celular é estimulada por uma variedade de sinais bioquímicos. Entretanto, outros sinais inibem a divisão celular. Em células normais, estes fatores contrários compensam uns aos outros com o resultado de que ocorre crescimento de modo regulado. Em células cancerosas, o crescimento é desregulado porque os sinais estimulantes prevale-

cem. Durante o progresso para malignidade, as células cancerosas perdem a capacidade de responder apropriadamente a sinais que inibem o crescimento. Por exemplo, as células em adenomas intestinais geralmente não respondem mais a  $TGF\beta$ , uma proteína que instrui pRB a bloquear a progressão do ciclo celular. Quando esse bloqueio falha, as células avançam de  $G_1$  para S, replicam seu DNA e dividem-se. Tais



células estão então no caminho para a formação de um tumor maligno.

3. *As células cancerosas podem escapar da morte celular programada.* Como vimos, p53 tem um papel principal em proteger o organismo do acúmulo de células danificadas que podem pôr em risco sua vida. Por meio de mecanismos ainda incompletamente compreendidos, p53 manda as células danificadas para uma via de autodestruição que as elimina do organismo. Quando p53 funciona mal, essa via de autodestruição é bloqueada, e as células danificadas sobrevivem e se multiplicam. Tais células provavelmente produzem descendentes que são ainda mais anormais que elas. Consequentemente, linhagens derivadas de células danificadas são propensas a avançar para um estado canceroso. A capacidade de escapar da morte celular programada é portanto uma característica chave na progressão para câncer maligno.
4. *As células cancerosas adquirem um potencial replicativo ilimitado.* Células normais são capazes de se dividir cerca de 60 a 70 vezes. Essa limitação surge da perda inexorável de DNA das pontas dos cromossomos toda vez que o DNA se replica (Cap. 10). O efeito cumulativo dessa perda cria uma capacidade reprodutiva finita em todas as linhagens celulares. Células que vão além do limite reprodutivo tornam-se geneticamente instáveis e morrem. As células cancerosas transcendem este limite repondo seu DNA perdido. Elas o fazem aumentando a atividade da enzima telomerase, que adiciona seqüências de DNA às pontas dos cromossomos. Quando as células adquiriram um potencial replicativo ilimitado superando o DNA perdido nas pontas dos cromossomos, diz-se que elas estão *imortalizadas*.
5. *As células cancerosas desenvolvem meios de se nutrir.* Qualquer tecido em um organismo multicelular complexo precisa de um sistema vascular para levar nutrientes para ele. Em humanos e outros animais vertebrados, o sistema circulatório fornece esta função. As células em tumores pré-malignos não crescem agressivamente porque não são diretamente alimentadas pelo sistema circulatório. Entretanto, quando vasos sanguíneos são induzidos a crescer entre tais células, por um processo chamado de **angiogênese**, o tumor é alimentado e pode ser ampliado. Assim, uma etapa fundamental na progressão do câncer maligno é a indução de crescimento de vasos sanguíneos pelas células do tumor. São conhecidos muitos fatores que induzem ou inibem a angiogênese. Em tecidos normais, esses fatores são mantidos balanceados de modo que vasos podem crescer apropriadamente no corpo. Em tecidos cancerosos, o equilíbrio pende para os fatores indutores, que atuam para estimular o desenvolvimento de vasos sanguíneos. Uma vez que capilares cresceram para um tumor, os meios de nutrição estão disponíveis. O tumor pode então alimentar-se e crescer a um tamanho em que se torna perigoso para o organismo.

6. *As células cancerosas adquirem a capacidade de invadir outros tecidos e colonizá-los.* Mais de 90% de todas as mortes por câncer são causadas por metástase do câncer para outras partes do corpo. Quando tumores metastatizam, as células cancerosas destacam-se do tumor primário e percorrem pela corrente sanguínea até outro local, onde estabelecem uma nova e longa relação, e no final letal, com as células vizinhas. Devem ocorrer mudanças profundas nas superfícies das células cancerosas para que este processo se instale. Quando isto ocorre, podem desenvolver-se tumores secundários em tecidos bem distantes do tumor primário. Cânceres que se espalharam deste modo são extremamente difíceis de controlar e erradicar. Metástase é, portanto, a ocorrência mais séria na progressão de um câncer.

Vários estudos estabeleceram que mutação somática é a base para desenvolvimento e progressão de todos os tipos de câncer. À medida que um câncer progride na via para malignidade, suas células tornam-se progressivamente desreguladas. Mutações acumulam-se, e cromossomos inteiros ou segmentos cromossômicos podem perder-se. Essa instabilidade genética aumenta a probabilidade de que o câncer desenvolva cada um dos marcos discutidos anteriormente.

Em razão da importância de mutações somáticas na etiologia do câncer, fatores que aumentam a taxa de mutação elevam a incidência de câncer. Hoje em dia, muitos países mantêm programas de pesquisa para identificar agentes mutagênicos e carcinogênicos (veja no Cap. 13 uma discussão do teste de Ames para identificação de mutágenos químicos). Quando tais agentes são identificados, autoridades de saúde pública criam políticas para diminuir exposição humana a eles. Entretanto, nenhum ambiente está livre de carcinógenos, e comportamentos humanos que contribuem para o risco de câncer tais como fumo, exposição excessiva ao sol e consumo de alimentos gordurosos que contêm pouca fibra são difíceis de mudar. A compreensão dos processos que causam câncer avançou significativamente durante a última década. Na próxima década, esperamos que esta compreensão resulte em estratégias mais efetivas para prevenção e tratamento do câncer.

#### ■ PONTOS IMPORTANTES

- Tipos diferentes de câncer estão associados a mutações em genes diferentes.
- As células cancerosas podem estimular seu próprio crescimento e divisão.
- As células cancerosas não respondem a fatores que inibem crescimento celular.
- As células cancerosas podem escapar dos mecanismos naturais que matam células anormais.
- As células cancerosas imortalizadas podem dividir-se indefinidamente.
- Tumores podem aumentar quando induzem o crescimento de vasos sanguíneos para nutrir suas células.
- As células cancerosas metastáticas podem invadir outros tecidos e colonizá-los.





## UM MARCO NA GENÉTICA: A Identificação do Gene *BRCA1*

Em 1994, um grupo de pesquisa do National Institute of Environmental Health Sciences in Research Triangle, Carolina do Norte, McGill University em Montreal, Canadá, a University of Utah em Salt Lake City, Myriad Genetics, uma empresa com base em Salt Lake City, e o Eli Lilly, o gigante farmacêutico com base em Indianapolis, Indiana, em um total de 45 pessoas, publicaram um artigo relatando a identificação de um gene envolvido no câncer hereditário de mama e ovário.<sup>1</sup>

Este gene era há muito procurado. No início dos anos 1980, evidências começaram a sugerir que um ou mais genes predispoem mulheres a desenvolver câncer de mama ou ovário. Em geral, esses cânceres aparecem cedo na vida adulta e costumam ser fatais. A análise de heredogramas mostrando a herança desses cânceres foi complicada por variação na idade de início, por penetrância incompleta e pela natureza limitada ao sexo da característica. Entretanto, em 1990 um dos genes que predispoem ao câncer foi mapeado no cromossomo 17. Progressos posteriores refinaram a localização deste gene, chamado de *BRCA1*, no meio do braço longo do cromossomo. Pesquisadores foram capazes de situar *BRCA1* dentro de uma região contendo cerca de 600 quilobases de DNA. Então, o problema foi identificar qual segmento da região com 600 kb era de fato o gene *BRCA1*.

Para abordar este problema, o consórcio de grupos de pesquisa coletou clones de DNA de dentro da região candidata. Foram obtidos todos os tipos de clones – clones grandes propagados em cromossomos artificiais de leveduras (YAC) e cromossomos artificiais de bactérias (BAC) e clones menores propagados em bacteriófagos e cosmídeos. A coleção de clones permitiu que pesquisadores construíssem um mapa físico da região candidata. Pontos ao longo deste mapa foram então analisados quanto a expressão de genes transcritos. Foram identificadas um total de 65 seqüências expressas. O DNA correspondente às partes codificantes destes transcritos foi analisado na esperança de se encontrar mutações em pessoas que supostamente tinham um alelo de predisposição do gene de câncer de mama. A evidência de várias hipotéticas portadoras indicou que *BRCA1* está no meio da região de 600 kb. Ele possui 22 éxons dispersos em 100 kb, e seu mRNA com 7,8 kb de tamanho codifica um polipeptídeo com 1.863 aminoácidos.

Os pesquisadores encontraram quatro mutações dentro da seqüência codificante candidata: (1) uma deleção de 11 pares de bases no éxon 2 que causa uma mudança na matriz de leitura, (2) uma mutação sem sentido no éxon 11 que causa término prematuro de

síntese de polipeptídeos, (3) uma inserção de 1 par de bases no éxon 20 que causa uma mudança de matriz de leitura e (4) uma mutação de sentido trocado no éxon 21 que substitui metionina por arginina na proteína. Além disso, eles encontraram um caso no qual o transcrito do gene candidato estava ausente, possivelmente porque uma mutação regulatória tinha ocorrido em algum elemento não codificante. Cada uma destas cinco mutações era de uma família diferente.

As portadoras hipotéticas de outras três famílias não apresentavam mutações nos genes candidatos a codificar a seqüência. Entretanto, nessas famílias, a evidência de ligação entre o gene de predisposição ao câncer e os marcadores moleculares no braço longo do cromossomo 17 era fraca. Assim, eles podem representar casos em que está envolvido um gene à exceção do *BRCA1*. De fato, no final desta publicação, os pesquisadores notaram que em uma dessas famílias a predisposição a desenvolver câncer de mama podia estar ligada a *BRCA2*, um gene em um cromossomo diferente. Alternativamente, essas três famílias poderiam ter mutações de predisposição ao câncer nos elementos não codificantes de *BRCA1*.

Desde 1994, muito mais mutações foram identificadas no gene *BRCA1*. A descoberta dessas mutações possibilitou o desenvolvimento de testes de DNA que ajudam a identificar mulheres que, devido a seus genótipos, estão em alto risco de desenvolver câncer de mama e de ovário. Estes testes representam as primeiras etapas no combate a esta doença fetal.

### QUESTÕES PARA DISCUSSÕES

1. O gene *BRCA1* foi identificado por um consórcio de pesquisadores que foram apoiados por uma combinação de fundos públicos e privados. Um dos principais pesquisadores, Mark Skolnick, era tanto professor da University of Utah, onde recebeu verbas governamentais, quanto funcionário de uma empresa de biotecnologia, Myriad Genetics, que subsequente desenvolveu testes comerciais para alelos *BRCA1* mutantes. Que questões éticas estão envolvidas quando uma pesquisa é apoiada conjuntamente por verbas governamentais e de empresas, especialmente quando a empresa visa lucros da pesquisa?
2. Uma mulher com teste positivo para um alelo mutante de predisposição do gene *BRCA1* tem alta probabilidade (>80%) de desenvolver câncer de mama ou ovário – e de morrer de um destes tipos de câncer. Que medidas estão disponíveis para evitar a ocorrência destes cânceres? Se você fosse uma destas mulheres, o que você faria?

<sup>1</sup>Miki, Y., et al. 1994. A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science* 266: 66-71.



## Exercícios Básicos

Ilustram a análise genética básica.

1. Que pontos de verificação do ciclo celular impedem que uma célula replique um DNA danificado?

**Resposta:** O ponto de verificação *START* em metade de  $G_1$  do ciclo celular.

2. (a) Em que classes de genes mutações dominantes de ganho de função causam câncer? (b) Em que classes de genes mutações recessivas de ganho de função causam câncer?

**Resposta:** (a) Oncogenes. (b) Genes supressores tumorais.

3. Por que alguns rearranjos cromossômicos causam câncer?

**Resposta:** Os pontos de quebra desses rearranjos em geral justapõem um oncogene celular a um promotor que estimula a expressão vigorosa do oncogene. A hiperexpressão do produto gênico pode resultar em divisão celular e crescimento excessivos.

4. Câncer intestinal ocorre em indivíduos com mutações inativadoras no gene *APC*. Explique como isto também pode ocorrer em pessoas com mutações no gene de  $\beta$ -catenina.

**Resposta:** Uma mutação que impede especificamente que  $\beta$ -catenina se ligue a pAPC pode causar câncer. A  $\beta$ -catenina que não pode ligar-se a pAPC estaria disponível para ligar-se aos fatores de transcrição que estimulam a expressão de genes cujos produtos promovem divisão celular e crescimento.

5. Que gene supressor tumoral está mais frequentemente mutado em cânceres humanos?

**Resposta:** *TP53*, o gene que codifica p53.

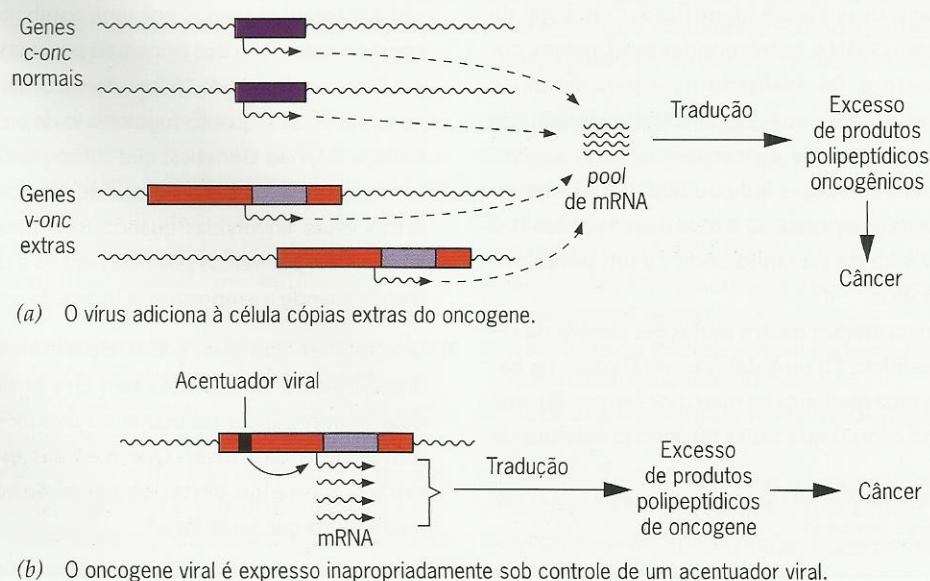
## Testar Seus Conhecimentos

Integra conceitos e técnicas diferentes.

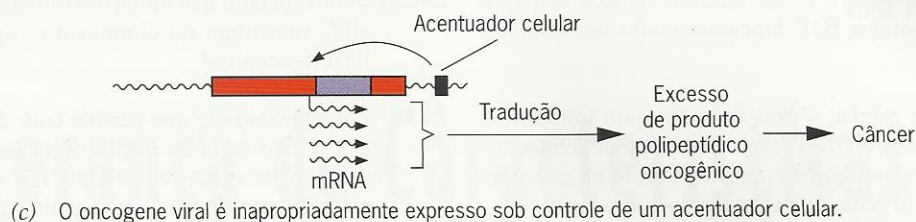
1. Um oncogene dentro do genoma de um retrovírus tem alta probabilidade de causar câncer, mas um oncogene em sua posição cromossômica normal não. Se estes dois oncogenes codificam exatamente o mesmo polipeptídeo, como explicamos suas propriedades diferentes?

**Resposta:** Existem pelo menos três possibilidades. Uma (a) é que o vírus simplesmente adiciona cópias extras do oncogene à célula, e coletivamente elas produzem muito deste polipeptídeo. Um excesso de polipeptídeo pode causar divisão celular descontrolada; isto é, câncer. Outra possibilidade (b) é que o oncogene viral é expresso inapropriadamente sob o controle de acentuadores no DNA viral. Esses acentuadores podem ati-

var o oncogene a ser expresso na época errada ou ser hiperexpresso constitutivamente. Em ambos os casos, o polipeptídeo seria inapropriadamente produzido e poderia assim perturbar os controles normais da divisão celular. Uma terceira possibilidade (c) é que a integração do vírus aos cromossomos da célula infectada pode colocar o oncogene viral na vizinhança de um acentuador no DNA cromossômico, e este acentuador pode provocar uma expressão inapropriada. Todas as três explicações destacam a idéia de que a expressão de um oncogene deve ser corretamente regulada. Má expressão ou hiperexpressão pode resultar em divisão celular descontrolada. Veja o diagrama seguinte.



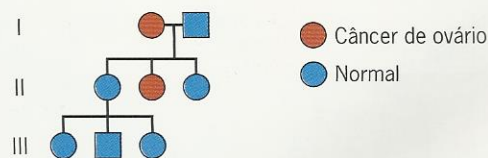




## Questões e Problemas

Acentuam a compreensão e desenvolvem as habilidades analíticas.

- 24.1. Muitos cânceres parecem envolver fatores ambientais. Por que então o câncer é chamado de uma doença genética?
- 24.2. Tanto células embrionárias quanto cancerosas dividem-se rapidamente. Como estes dois tipos de células podem ser distinguidos um do outro?
- 24.3. A maioria das células cancerosas são aneuplóides. Sugira como a aneuploidia pode contribuir para desregulação do ciclo celular.
- 24.4. Você esperaria encontrar um retrovírus indutor de tumor que leve um gene supressor tumoral celular processado em seu genoma?
- 24.5. Como sabemos que oncogenes celulares normais não são simplesmente oncogenes retrovirais integrados que adquiriram a regulação apropriada?
- 24.6. Como pode a ausência de íntrons em um oncogene retroviral explicar hiperexpressão do gene nos tecidos de um animal infectado?
- 24.7. Quando oncogenes celulares são isolados de animais diferentes e comparados, as seqüências de aminoácidos dos polipeptídeos que eles codificam são muito similares. O que isto sugere sobre as funções desses polipeptídeos?
- 24.8. A maioria dos oncogenes *c-ras* obtidos de tecidos cancerosos têm mutações no códon 12, 59 ou 61 na seqüência codificante. Sugira uma explicação.
- 24.9. Quando um oncogene mutante *c-H-ras* com uma substituição de glicina por valina no códon 12 é transfectado para células cultivadas NIH 3T3, ele transforma essas células em células cancerosas. Quando o mesmo oncogene mutante é transfectado para células embrionárias cultivadas, ele não as transforma. Por quê?
- 24.10. Uma mutação no oncogene celular *ras* pode causar câncer quando na condição heterozigota, mas uma mutação no gene supressor tumoral *RB* só pode causar câncer quando na condição homozigota. O que esta diferença entre mutações dominantes e recessivas significa sobre os papéis dos produtos gênicos de *ras* e *RB* em atividades celulares normais?
- 24.11. Explique por que indivíduos que desenvolvem retinoblastoma não hereditário têm tumores em apenas um olho, enquanto indivíduos com retinoblastoma hereditário geralmente desenvolvem tumores em ambos os olhos.
- 24.12. Cânceres hereditários como retinoblastoma apresentam um padrão dominante de herança. Entretanto, o defeito genético subjacente é uma mutação recessiva de perda de função, em geral o resultado de uma deleção. Como o padrão dominante de herança pode ser conciliado com a natureza recessiva da mutação?
- 24.13. Aproximadamente 5% das pessoas que herdam um gene *RB* inativo não desenvolvem retinoblastoma. Use esta estatística para estimar o número de divisões celulares que formam o tecido retiniano do olho. Suponha que a taxa em que as mutações somáticas inativam o gene *RB* é de 1 mutação por  $10^6$  divisões celulares.
- 24.14. O heredograma abaixo mostra a herança de câncer ovariano familiar causado por uma mutação no gene *BRC1*. II-1 deveria ser testada quanto à presença da mutação predisponente? Comente as vantagens e desvantagens da investigação.



- 24.15. Em que sentido pRB é um regulador negativo de fatores de transcrição E2F?
- 24.16. Um fator de transcrição E2F específico reconhece a seqüência TTTCGCGC no promotor de seu gene alvo. Uma mutação sensível a temperatura no gene codificante deste fator de transcrição E2F altera a capacidade de seu produto protéico de ativar a transcrição; a 25°C, a proteína mutante ativa normalmente a transcrição, mas a 35°C ela não ativa a transcrição. Entretanto, a capacidade da proteína de reconhecer sua seqüência de DNA alvo não é prejudicada em nenhuma destas temperaturas. Seria esperado que células heterozigotas para esta mutação sensível a temperatura se



- dividam normalmente a 25°C? a 35°C? As suas respostas mudam se a proteína E2F funcionar como um homodímero?
- 24.17.** Durante o ciclo celular, a proteína p16 é um inibidor de atividade de ciclina/CDK. Preveja o fenótipo de células homozigotas para a mutação de perda de função no gene que codifica p16. Este gene seria classificado como um proto-oncogene ou como um gene supressor tumoral?
- 24.18.** O gene *BCL-2* codifica uma proteína que reprime a via para morte celular programada. Preveja o fenótipo de células heterozigotas para uma mutação ativadora dominante neste gene. O gene *BCL-2* seria classificado como um proto-oncogene ou como um gene supressor tumoral?
- 24.19.** O produto protéico do gene *BAX* regula negativamente o produto protéico do gene *BCL-2*, isto é, a proteína BAX interfere na função da proteína BCL-2. Preveja o fenótipo de células homozigotas para uma mutação de perda de função no gene *BAX*. Este gene seria classificado como um proto-oncogene ou como um gene supressor tumoral?
- 24.20.** As células cancerosas frequentemente são homozigotas para mutações de perda de função no gene *TP53*, e muitas dessas mutações estão mapeadas na parte do *TP53* que codifica o domínio de ligação ao DNA de p53. Explique como tais mutações contribuem para o fenótipo de câncer das células.
- 24.21.** Suponha que uma célula é heterozigota para uma mutação que faz com que p53 se ligue fortemente e constitutivamente ao DNA de seus genes alvos. Como esta mutação afeta o ciclo celular? Seria esperado que tal célula fosse mais ou menos sensível aos efeitos de radiação ionizante?
- 24.22.** Camundongos homozigotos para uma mutação nocaute do gene *TP53* são viáveis. Seria esperado que eles sejam mais ou menos sensíveis aos efeitos letais de radiação ionizante?
- 24.23.** Seria esperado que mutações causadoras de câncer do gene *APC* aumentem ou diminuam a capacidade de pAPC de ligar  $\beta$ -catenina?
- 24.24.** Foi demonstrado que pessoas com dietas pobres em fibras e ricas em comidas gordurosas têm risco aumentado de desenvolver câncer colorretal. Tais dietas podem irritar o revestimento epitelial do intestino grosso. Como tal irritação pode contribuir para o risco aumentado de câncer colorretal?
- 24.25.** Camundongos que são heterozigotos para uma mutação nocaute no gene *RB* desenvolvem tumores na hipófise e na tireóide. Camundongos que são homozigotos para esta mutação morrem durante o desenvolvimento embrionário. Camundongos que são homozigotos para uma mutação nocaute no gene que codifica o homólogo p130 de RB e heterozigotos para uma mutação nocaute no gene que codifica o homólogo p107 de RB não têm tendência a desenvolver tumores. Entretanto, os homozigotos para mutações nocaute em ambos estes genes morrem durante o desenvolvimento embrionário. O que estes achados sugerem quanto aos papéis dos genes *RB*, *p139* e *p107* em embriões e adultos?
- 24.26.** O RNA mensageiro do gene *KAI1* é fortemente expresso em tecidos prostáticos normais mas fracamente expresso em linhagens celulares derivadas de cânceres prostáticos metastáticos. O que este achado sugere sobre o papel do produto do gene KAI1 na etiologia do câncer de próstata?
- 24.27.** A proteína p21 é fortemente expressa em células que foram irradiadas. Pesquisadores pensaram que essa forte expressão era provocada por ativação transcrricional do gene *p21* pela proteína p53 agindo como um fator transcrricional. Esta hipótese ajusta-se às observações de que a expressão de p21 é induzida por tratamento com radiação em camundongos homozigotos para uma mutação nocaute no gene *TP53*? Explique.



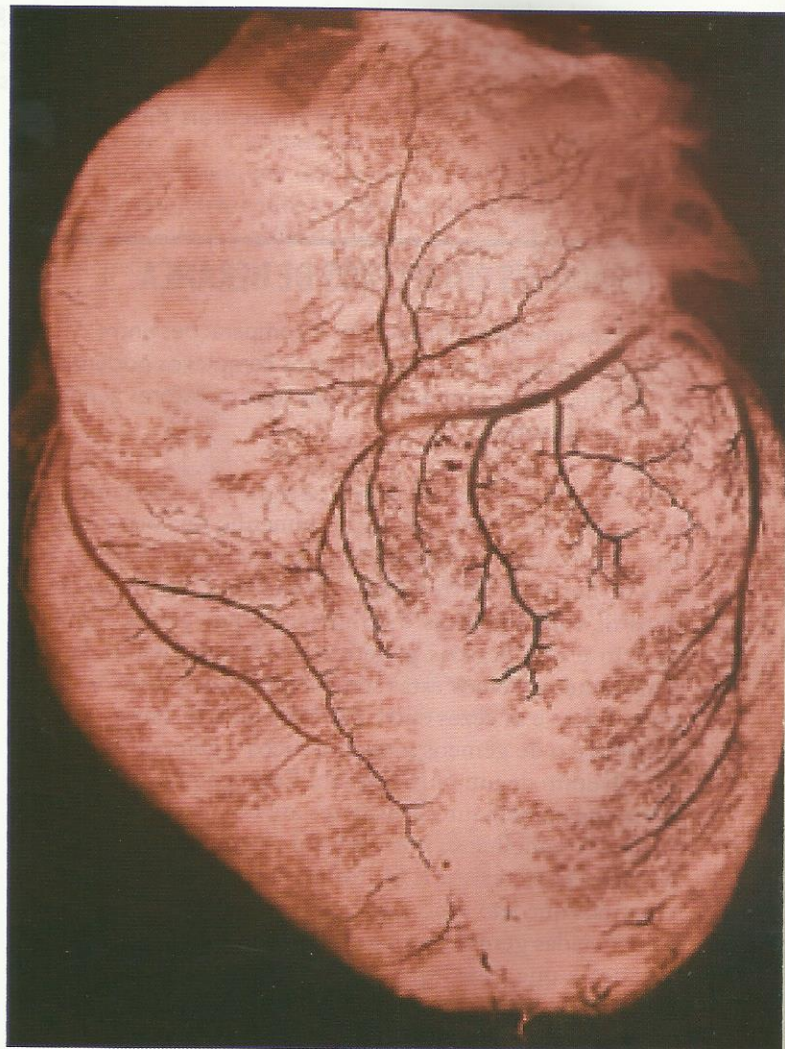
# Herança de Características Complexas

## Tópicos

- Padrões Complexos de Herança
- Estatísticas de Genética Quantitativa
- Análise de Características Quantitativas
- Análise de Endogamia e Co-Ancestralidade
- Correlações Entre Parentes
- Genética Quantitativa de Características Comportamentais Humanas

## Doença Cardiovascular: Uma Combinação de Fatores Genéticos e Ambientais

No final de dezembro de 1996, Paul Reston, um professor de Biologia de 47 anos de idade de um colégio do subúrbio de Pittsburgh, Pennsylvania, estava aplicando exames em uma manhã de sábado. Ele estava um pouco cansado naquele dia e sofria distúrbios estomacais. Também tinha uma leve dor no braço esquerdo e no ombro. Estes sintomas persistiram por alguns dias. No começo, Mr. Reston pensou que tinha um caso brando de gripe, mas a dor no braço e no ombro sugeriram outra possibilidade: ele estava tendo um ataque cardíaco. Esta possibilidade pareceu mais real quando ele se lembrou que seu pai havia morrido de um súbito ataque cardíaco muitos anos antes na idade relativamente jovem de 45 anos. Após uma conversa telefônica com uma enfermeira de sua clínica de saúde, Mr. Reston pediu que seu filho o levasse até o hospital próximo, onde ele ficou duas horas na sala de emergência. O médico que atendeu Mr. Reston fez uma bateria de testes para avaliar



Um coração humano mostrando as artérias e veias que fornecem sangue aos músculos cardíacos.

sua condição. Seu batimento cardíaco era regular, sua pressão arterial era normal e o eletrocardiograma não revelou anomalias. Os testes bioquímicos para sinais indicativos de danos cardíacos também foram negati-



vos. Além disso, à exceção de história familiar de doença cardíaca, Mr. Reston não apresentava outros fatores de risco importantes. Ele não estava acima do peso, não fumava e fazia exercícios regularmente. O médico liberou Mr. Reston mas o alertou que voltasse ao hospital na semana seguinte para um teste de esforço cardíaco. Na segunda feira seguinte, ele foi testado quanto ao funcionamento cardíaco correndo em uma esteira. Os resultados do teste foram bons. Com base em seu desempenho, o cardiologista supervisor concluiu que Mr. Reston tinha menos de 1% de chance de sofrer um ataque cardíaco fatal.

A despeito de sua história familiar de doença cardíaca, o risco de Mr. Reston desenvolver esta doença era baixo. O cardiologista explicou que a doença cardíaca era uma característica complexa influenciada por muitos fatores: dieta, atividade física e tabagismo, por exemplo, assim como por um número razoavelmente grande de genes. Como o pai de Mr. Reston havia morrido de um infarto do miocárdio, Mr. Reston podia ter herdado genes que o colocavam em risco. Entretanto, o cardiologista destacou que a doença cardíaca não é herdada como uma característica mendeliana simples; em vez disso, ela envolve uma interação de muitos fatores genéticos e ambientais diferentes.

## PADRÕES COMPLEXOS DE HERANÇA

Experimentos reprodutivos e comparações entre parentes revelam que fenótipos complexos podem ser influenciados por uma combinação de fatores genéticos e ambientais.

Muitas características interessantes e importantes, tais como suscetibilidade a doenças, tamanho do corpo e vários aspectos comportamentais, não apresentam padrões simples de herança. Entretanto, sabemos que a variação nestes tipos de características é influenciada por genes. Uma indicação é que pessoas geneticamente aparentadas exibem semelhanças. Vemos estas semelhanças entre irmãos, entre pais e filhos e às vezes entre parentes mais distantes. O caso extremo é o de gêmeos monozigóticos – gêmeos que se desenvolveram de um único zigoto. Esses gêmeos geralmente são muito semelhantes e às vezes até compartilham maneirismos. A semelhança entre parentes é uma forte indicação de que as características anatômicas, fisiológicas e comportamentais que não mostram padrões simples de herança são influenciadas geneticamente. Outra indicação de influência genética é que estes tipos de características podem ser alterados por cruzamento seletivo. Na agricultura, cultivos e criações foram moldados pela propagação de indivíduos com características desejáveis – maior conteúdo protéico, gordura corpórea reduzida, maior produtividade, resistência a doenças e assim em diante. Tal capacidade de mudar fenótipos por cruzamento seletivo indica que a variação nas características tem uma base genética. Entretanto, sabemos por experimentação que a base genética é complexa. Nas seções seguintes, consideramos dois tipos de características geneticamente complexas: (1) características que variam continuamente em uma população e (2) características que

se manifestam quando é ultrapassado um limiar genético subjacente.

## Características Quantitativas

A natureza está cheia de exemplos de características que variam continuamente – tamanho do corpo, altura, peso, a atividade de uma determinada enzima, capacidade reprodutiva e vigor geral e adaptação, para citar apenas algumas. Estas estão entre as mais interessantes características em biologia, especialmente do ponto de vista evolutivo. Um fenótipo parece misturar-se imperceptivelmente com o seguinte. Entretanto, se examinarmos muitos indivíduos em uma população, geralmente encontramos diferenças significativas entre eles. Em geral, esse tipo de variação pode ser quantificado medindo-se a característica em uma amostra de indivíduos da população. Podemos, por exemplo, capturar camundongos em um celeiro e pesá-los ou coletar espigas de milho em um campo e contar o número de grãos em cada uma. Com tal enfoque quantitativo, o fenótipo de cada indivíduo na amostra é reduzido a um número. Esses números podem ser analisados com uma variedade de técnicas estatísticas, permitindo-nos estudar a característica e, em particular, investigar sua base genética. As características que são passíveis deste tipo de tratamento são chamadas de **características quantitativas**. Sua particularidade essencial é que elas podem ser medidas.

Charles Darwin pensou que tal variação quantitativa era crucial para a evolução de uma espécie. Ele formulou a hipótese de que esta variação fornecia a matéria prima para uma espécie mudar gradualmente com o tempo. O mecanismo para a mudança que Darwin propôs foi a reprodução diferencial de indivíduos de acordo com o fenótipo – o processo que ele denominou seleção natural. Por exemplo, os



maiores indivíduos em uma espécie poderiam produzir mais prole que os outros indivíduos. Considerando que o tamanho é herdado, devemos esperar que a geração seguinte seja um pouco maior, em média, que a geração parental. Com o tempo, as mudanças de tamanho iriam acumulando-se, e os membros da espécie seriam consistentemente maiores que seus ancestrais foram.

Quando Darwin propôs suas idéias em 1859, não havia compreensão do mecanismo da hereditariedade. Os genes, os cromossomos e os princípios de segregação e distribuição independente ainda não tinham sido descobertos. Mendel foi contemporâneo de Darwin, mas sua pesquisa sobre o mecanismo de herança aparentemente nunca chamou a atenção de Darwin. Depois que as descobertas de Mendel foram reveladas em 1900, biólogos procuram intensamente ligá-las à teoria evolutiva de Darwin. A dificuldade foi que as características consideradas importantes para a evolução – características quantitativas tais como tamanho do corpo, vigor e capacidade reprodutiva – não eram “mendelianas”, isto é, não segregavam como fatores mendelianos simples. Em vez disso, sua herança era complexa. Por quase 20 anos, biólogos lutaram para conciliar a complexidade de características quantitativas com a simplicidade da teoria de Mendel. Alguns cientistas, tais como o estatístico britânico Karl Pearson, sustentavam que elas não podiam ser conciliadas. Esses cientistas diziam que os princípios de Mendel simplesmente não se aplicavam à análise de padrões complexos de herança. Entretanto, este ponto de vista acabou por ser modificado. Um volume crescente de evidências sugeriram que a herança complexa de características quantitativas podia ser explicada pela teoria mendeliana. Como veremos, dois *insights* possibilitaram isto:

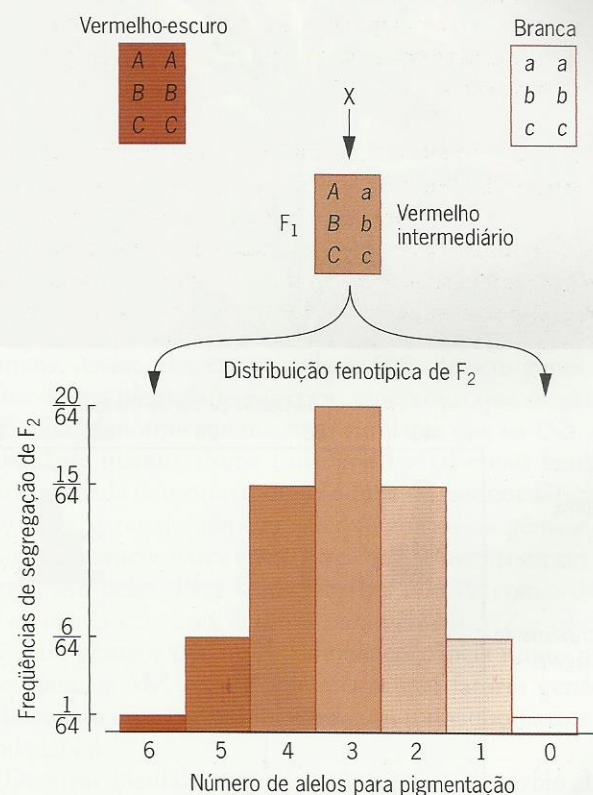
1. As características quantitativas são influenciadas por muitos genes.
2. As características quantitativas são influenciadas por muitos fatores ambientais.

É a combinação de influências genéticas e ambientais que torna difícil estudar as características quantitativas. Tais características dependem de uma mistura complexa de fatores; isto é, elas são **multifatoriais**.

O biólogo dinamarquês Wilhelm Johannsen foi uma das primeiras pessoas a mostrar que a variação em uma característica quantitativa se deve a fatores genéticos e ambientais. Johannsen estudou o peso das sementes da fava *Phaseolus vulgaris*. Entre as plantas disponíveis para ele, o peso das sementes variava de 150 mg a 900 mg. Johannsen estabeleceu linhagens de sementes nesta faixa e manteve cada linhagem por autofecundação durante várias gerações. As sementes de cada uma dessas linhagens “puras” tendiam a se parecer com a semente da qual se originaram. Essa capacidade de estabelecer linhagens de favas com pesos de sementes caracteristicamente diferentes indicou que alguma variação em tal característica decorre de diferenças genéticas. Entretanto, Johannsen observou que o peso da semente também variava

dentro de cada linhagem pura. Essa variação residual provavelmente não resultava de diferenças genéticas, pois cada linhagem tinha sido sistematicamente endocruzada com o objetivo de torná-la homozigota para seus genes. Isto devia ser consequência de variação em fatores não controlados no ambiente. O trabalho de Johannsen, publicado em 1903 e 1909, levou, portanto, à percepção de que a variação fenotípica em uma característica quantitativa tinha dois componentes – um genético e outro ambiental.

Outro escandinavo, Herman Nilsson-Ehle, forneceu evidências de que o componente genético dessa variação podia envolver as contribuições de vários genes diferentes. Nilsson-Ehle estudou a variação de cor em grãos de trigo. Quando ele cruzou uma variedade de grãos brancos com uma variedade de grãos vermelho-escuros, ele obteve uma  $F_1$  com um fenótipo vermelho intermediário (Fig. 25.1). A autofecundação da  $F_1$  produziu uma  $F_2$  com sete classes distintas, variando de branco a vermelho-escuro. O número de classes de  $F_2$  e a proporção fenotípica que Nilsson-Ehle observou sugeriram que três genes de distribuição independente estavam envolvidos na determinação da cor do grão. Nilsson-Ehle formulou a hipótese de que cada gene tem dois alelos, um responsável pela cor vermelha do grão e o



**Fig. 25.1** ■ Herança da cor de grãos em trigo. Três genes que se distribuem independentemente (A, B e C) supostamente controlam a cor dos grãos. Cada gene tem dois alelos. Os alelos para pigmentação (representados por letras maiúsculas) são semidominantes em relação aos alelos para falta de pigmentação (representados por letras minúsculas).



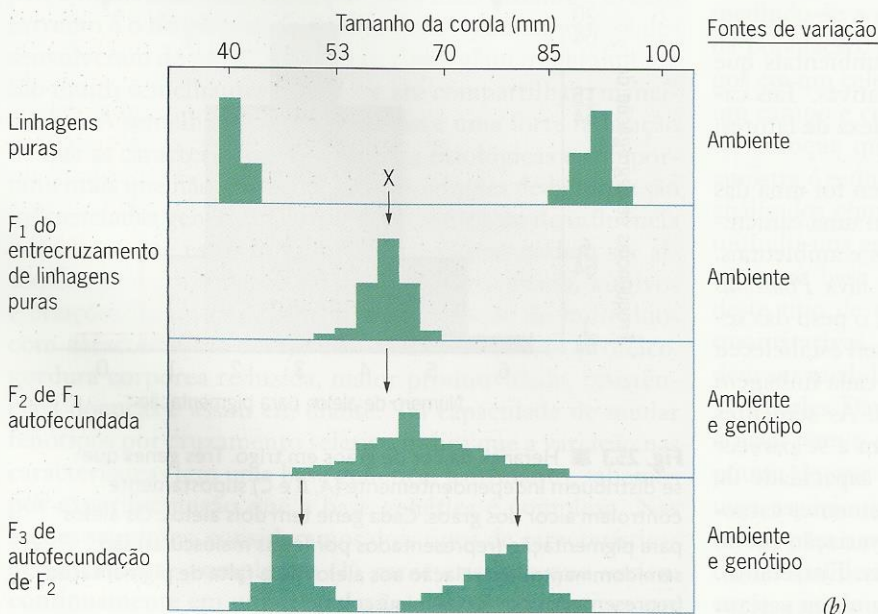
outro pela cor branca, e que o alelo para o grão de cor vermelha era semidominante em relação ao alelo para grão de cor branca. Com base nesta hipótese, o genótipo do genitor de grão branco podia ser representado por  $aa\ bb\ cc$  e o genótipo do genitor de grão vermelho podia ser representado como  $AA\ BB\ CC$ . O genótipo de  $F_1$  podia ser  $Aa\ Bb\ Cc$  e a  $F_2$  podia conter uma gama de genótipos que iriam diferir no número de alelos semidominantes, isto é, contribuintes de pigmento, presentes. Cada classe de fenótipo na  $F_2$  levaria um número diferente de alelos contribuintes de pigmento. A classe branca, por exemplo, levaria zero de tais alelos, a classe de vermelho intermediário levaria três e a classe de vermelho-escuro levaria seis. O trabalho de Nilsson-Ehle, publicado em 1909, mostrou que um padrão complexo de herança podia ser explicado pela segregação e distribuição de vários genes.

O geneticista americano Edward M. East ampliou os estudos de Nilsson-Ehle para uma característica que não apre-

senta proporções mendelianas simples na  $F_2$ . East estudou o tamanho da corola em flores do tabaco (Fig. 25.2a). Em uma linhagem pura, o tamanho da corola tinha em média 41 mm; em outra, tinha média de 93 mm. Dentro de cada linhagem pura, East observou alguma variação fenotípica, supostamente o resultado de influências ambientais (Fig. 25.2b). Cruzando as duas linhagens, East obteve uma  $F_1$  que tinha um tamanho de corola intermediário e aproximadamente a mesma quantidade de variação que ele viu dentro de cada uma das linhagens parentais. Quando East entrecruzou plantas de  $F_1$ , ele obteve uma  $F_2$  com corolas mais ou menos do mesmo tamanho, em média, que ele viu na  $F_1$ . Entretanto, as plantas de  $F_2$  eram muito mais variáveis que as de  $F_1$ . Essa variabilidade decorria de duas fontes: (1) a segregação e a distribuição independente de pares de diferentes alelos que controlam o tamanho da corola e (2) fatores ambientais. East endocruzou algumas plantas de  $F_2$  para produzir uma  $F_3$  e observou menos variação dentro das diferentes linhagens de



(a)



(b)

**Fig. 25.2** ■ (a) Flores de tabaco mostrando a corola longa. (b) Herança de tamanho da corola em tabaco. Pelo menos cinco genes parecem estar envolvidos.

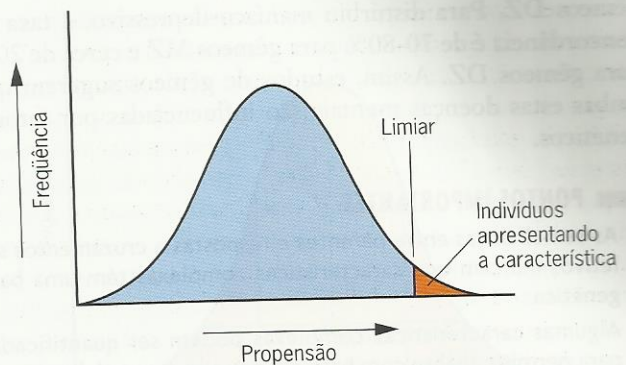


F<sub>3</sub> do que na F<sub>2</sub>. A quantidade reduzida de variação dentro das linhagens F<sub>3</sub> supostamente resultava da segregação de menos diferenças alélicas. Assim, o padrão complexo de herança que East observou no tamanho da corola podia ser explicado por uma combinação de segregação genética e influências ambientais.

Quanto genes estavam envolvidos na determinação do tamanho da corola nas linhagens de tabaco de East? Podemos fazer uma suposição grosseira comparando as plantas de F<sub>2</sub> com cada uma das linhagens parentais endocruzadas. Suponhamos que a linhagem com as corolas mais curtas fosse homozigota para um grupo de alelos e que a linhagem com as corolas maiores fosse homozigota para outro grupo de alelos. Além disso, suponhamos que os alelos com corolas longas são semidominantes, que todos os genes controladores de tamanho se distribuam independentemente e que cada gene faça uma contribuição igual para o fenótipo. Se o tamanho da corola fosse determinado por um gene, com os alelos *a* (para corola pequena) e *A* (para corola longa), esperaríamos que 1/4 das plantas F<sub>2</sub> tivessem corolas pequenas (como a linhagem parental pequena) e 1/4 tivessem corolas longas (como a linhagem parental longa). Se dois genes determinassem o tamanho da corola, esperaríamos 1/16 das plantas F<sub>2</sub> assemelhando-se ao genitor com corola pequena e 1/16 assemelhando-se ao genitor com corola longa. Se três genes estivessem envolvidos, a frequência de cada tipo parental na F<sub>2</sub> seria de 1/64 e, se quatro genes estivessem envolvidos, ela seria de 1/256. Com cinco genes, as frequências parentais na F<sub>2</sub> seriam cada uma de 1/1.024. East estudou 444 plantas de F<sub>2</sub> e falhou em encontrar até mesmo uma com um dos fenótipos parentais. Essa falha pareceria excluir a hipótese de quatro ou menos genes controlando o tamanho da corola. Assim, podemos concluir que pelo menos cinco genes são responsáveis pela diferença no tamanho da corola entre duas linhagens endocruzadas de East.

### Características com Limiar

As características de variação contínua tais como tamanho da corola, cor do grão e tamanho da fava são controladas por vários fatores, tanto genéticos quanto ambientais. Geneticistas descobriram que algumas características descontínuas também parecem ser multifatoriais. Por exemplo, alguns seres humanos nascem com fenda labial, uma condição causada por um erro no desenvolvimento embrionário. Embora estudos familiares tenham sugerido que a fenda labial é consequência de fatores genéticos e ambientais, ela não é uma característica quantitativa no sentido usual. Pessoas a expressam ou não. Entretanto, os fatores genéticos e ambientais que predispoem uma pessoa a apresentá-la parecem variar quantitativamente. Assim, embora a fenda labial se manifeste descontínuamente, sua variação subjacente parece ser contínua. Os geneticistas têm uma teoria de que, quando a variável subjacente excede um determinado nível, a característica aparece. Este tipo de característica é portanto cha-



**Fig. 25.3** ■ Um modelo para expressão de uma característica com limiar. Quando a variável subjacente, a “propensão”, atinge um valor de limiar, a característica é expressa. Supõe-se que essa variável é distribuída continuamente na população.

mado de **característica com limiar** (Fig. 25.3). A variável subjacente é uma função de fatores no genótipo bem como fatores no ambiente. Algumas pessoas possuem alelos que as predispoem a desenvolver fenda labial. Há, de fato, uma “tendência” genética a desenvolver a condição. Entretanto, apenas pessoas cuja tendência ultrapassa o limiar de fato estão em risco.

Em humanos, a evidência de que características tais como fenda labial são influenciadas por fatores genéticos vem de comparações entre parentes, especialmente gêmeos. Ocasionalmente, um zigoto humano divide-se e forma dois zigotos geneticamente idênticos. Os indivíduos que se desenvolvem destes zigotos são referidos como gêmeos **monozigóticos (MZ)**. Eles compartilham 100% de seus genes.<sup>1</sup> Mais frequentemente, dois ovócitos fertilizados independentemente se desenvolvem ao mesmo tempo no útero materno. Esses gêmeos **dizigóticos (DZ)** são tão parecidos quanto irmãos comuns. Assim, eles compartilham 50% de seus genes. Por causa de sua identidade genética, esperamos que os gêmeos MZ sejam fenotipicamente mais similares que os DZ. A similaridade quanto a uma característica tal como fenda labial é avaliada determinando-se a **taxa de concordância** – a fração de pares de gêmeos em que ambos os gêmeos exibem uma característica entre pares que se manifesta em pelo menos um deles. Para fenda labial, a taxa de concordância foi estimada em cerca de 40% para gêmeos MZ e cerca de 4% para gêmeos DZ. A taxa de concordância muito maior para gêmeos MZ sugere fortemente que fatores genéticos influenciam uma probabilidade de um indivíduo nascer com fenda labial.

Doenças mentais como esquizofrenia e distúrbio maníaco-depressivo também podem ser vistas como características com limiar. Para esquizofrenia, a taxa de concordância varia de 30 a 60% para gêmeos MZ e de 6 a 18% para

<sup>1</sup>N.T.: No início, porque depois, em consequência de mutações, começam a diferir e não são mais 100% iguais.



gêmeos DZ. Para distúrbio maniaco-depressivo, a taxa de concordância é de 70-80% para gêmeos MZ e cerca de 20% para gêmeos DZ. Assim, estudos de gêmeos sugerem que ambas estas doenças mentais são influenciadas por fatores genéticos.

### PONTOS IMPORTANTES

- As semelhanças entre parentes e respostas a cruzamentos seletivos indicam que características complexas têm uma base genética.
- Algumas características complexas podem ser quantificadas para permitir análise com base em princípios mendelianos.
- Muitos fatores genéticos e ambientais influenciam a variação observada em características quantitativas.
- Segregações fenotípicas podem fornecer um meio para estimar o número de genes que influenciam uma característica quantitativa.
- Características que se manifestam quando uma variável contínua subjacente (a "propensão") atinge um valor limiar podem ser influenciadas por fatores genéticos.
- Em humanos, a evidência de que uma característica limiar tem uma base genética vem de estudos de gêmeos.
- A taxa de concordância é a fração de pares de gêmeos em que ambos os gêmeos exibem uma característica entre pares que se manifesta em pelo menos um deles.

## ESTATÍSTICAS DE GENÉTICA QUANTITATIVA

As distribuições de frequências de características quantitativas podem ser caracterizadas por estatísticas simples.

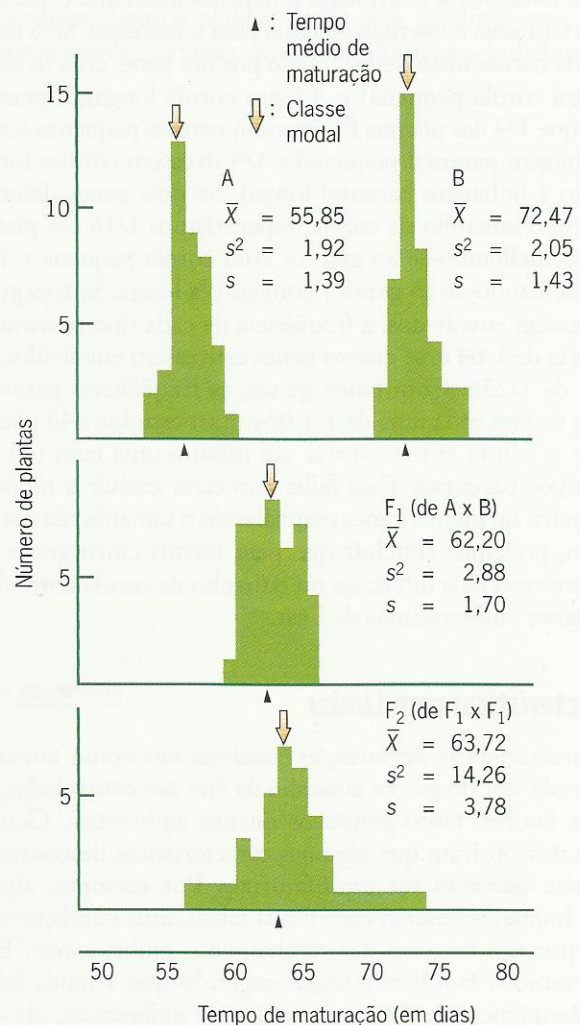
O marco de características quantitativas é que elas variam continuamente em uma população de indivíduos. Este tipo de variação cria um problema grande para o geneticista. As proporções de segregação são difíceis, ou mesmo impossíveis, de discernir porque o número de fenótipos é grande, e um fenótipo mistura-se imperceptivelmente com o seguinte. Para características de variação quantitativa, análises genéticas rotineiras do tipo que fizemos com a cor das sementes em ervilhas e com distúrbios humanos tais como albinismo estão fora de questão. Para estes tipos de características, devemos recorrer a um modelo diferente de análise, aquela que seja baseada em descrições estatísticas do fenótipo em uma população. Nas seções que se seguem, introduzimos os conceitos estatísticos básicos que são necessários para este tipo de análise.

### Distribuições de Frequência

A primeira etapa no estudo de qualquer característica quantitativa é coletar as medidas da característica de pessoas em uma população. Geralmente, apenas pequena fração de to-

dos os indivíduos na população pode ser medida. Nós chamamos este grupo de **amostra**. Os dados da amostra podem ser apresentados graficamente como uma **distribuição de frequência**. No gráfico, o eixo horizontal, ou  $x$ , mede os valores da característica. Este eixo é dividido em intervalos regulares que permitem classificar cada indivíduo na população quanto à característica. Assim, cada observação na amostra pode ser colocada em um dos intervalos do eixo  $x$ . O eixo vertical, ou  $y$ , mede a frequência das observações dentro de cada intervalo.

A Figura 25.4 mostra as distribuições de frequência que foram obtidas em um estudo genético de trigo. Os pesquisadores mediram o tempo que o trigo leva para amadurecer. Quatro populações diferentes de trigo foram cultivadas em



**Fig. 25.4** ■ Distribuições de frequência e estatística descritiva de tempo de maturidade em quatro populações de trigo. A e B são linhagens endogâmicas que foram cruzadas para produzir híbridos F<sub>1</sub>. Plantas F<sub>1</sub> foram entrecruzadas para produzir uma F<sub>2</sub>. Sementes de todas as quatro populações foram plantadas na mesma estação para determinar o tempo de maturidade. Em cada caso, os dados foram obtidos de 40 plantas. São citadas a média ( $\bar{X}$ ), a moda (seta vazada), a variância ( $s^2$ ) e o desvio padrão ( $s$ ).



um teste na mesma estação, e 40 plantas de cada população foram monitoradas até que o topo dos grãos amadurecessem. O tempo de maturação de cada planta foi registrado em dias. Duas das populações (A e B) eram linhagens endocruzadas e uma era uma  $F_1$  produzida por cruzamento destas duas linhagens. A quarta população era uma  $F_2$  produzida por cruzamento de plantas de  $F_1$ .

As duas linhagens parentais A e B eram variedades altamente endocruzadas que eram totalmente, ou quase totalmente, homozigotas. Como indicam as distribuições de frequência, a linhagem A amadureceu rapidamente e a linhagem B lentamente. A falta de superposição fenotípica entre as amostras destas duas linhagens demonstra sua distinção genética. Aparentemente, as linhagens A e B eram homozigotas para alelos diferentes de genes que controlam o tempo de maturação. Dentro de cada linhagem, entretanto, ainda havia alguma variação fenotípica, supostamente o resultado de diferenças microambientais dentro do teste.

As distribuições das amostras de  $F_1$  e  $F_2$  indicam que estas populações tinham tempos de maturação intermediários. Sua posição intermediária no eixo  $x$  sugere que os alelos controladores do tempo de maturação são semidominantes. Observe que a distribuição da amostra de  $F_2$  é consideravelmente mais larga que a de  $F_1$ . A variabilidade adicional vista na população  $F_2$  reflete a segregação genética que ocorre quando as plantas  $F_1$  são reproduzidas. Exploraremos agora os modos pelos quais geneticistas quantitativos resumem os dados em uma distribuição de frequência.

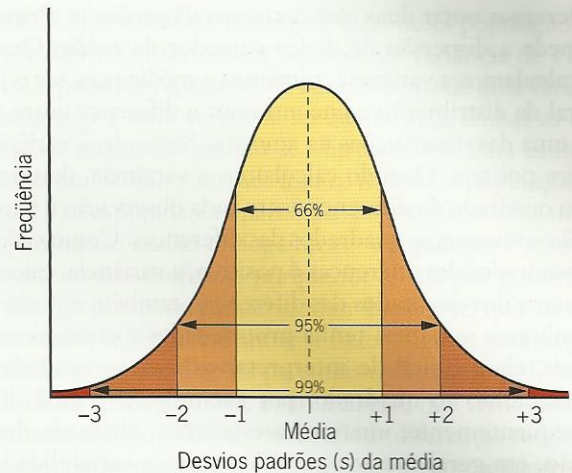
### As Classes Média e Modal

As características essenciais de uma distribuição de frequência podem ser resumidas por estatísticas simples calculadas a partir dos dados. Uma dessas estatísticas simples é chamada **média**. Ela nos dá o “centro” da distribuição – o valor “típico”. Nós calculamos a média da amostra ( $\bar{X}$ ) somando todos os dados na amostra e dividindo pelo número total de observações ( $n$ ). Em notação matemática, a média é:

$$\bar{X} = (\Sigma X_k) / n$$

A letra grega  $\Sigma$  nesta fórmula é uma abreviação matemática da soma de todas as medidas individuais na amostra; logo,  $\Sigma X_k = (X_1 + X_2 + X_3 + \dots + X_n)$ , onde  $X_k$  representa a  $k^{\text{a}}$  das  $n$  observações individuais. Na Figura 25.4, as posições das médias das amostras são indicadas por triângulos abaixo das distribuições; os valores numéricos dessas médias são dados à direita. As médias das amostras  $F_1$  e  $F_2$  são 62,20 e 63,72 dias, respectivamente. Ambas são um pouco menos que a média das médias das duas populações parentais endocruzadas (64,16 dias).

A **classe modal** em uma amostra é a classe que contém a maioria das observações. Como a média, ela também capta o “centro” da distribuição. Na Figura 25.4, as classes modais são indicadas por setas curtas. Vemos que em cada uma das



**Fig. 25.5** ■ Uma distribuição de frequência normal mostrando a porcentagem de medidas dentro de 1, 2 e 3 desvios padrões da média.

distribuições a média está dentro ou bem perto da classe modal. Esta coincidência reflete a simetria das distribuições. Em cada caso, números mais ou menos iguais de observações estão acima e abaixo da média e da classe modal. Nem todas as distribuições têm esta característica. Algumas estão desviadas, com a maioria das observações agrupadas em uma extremidade e apenas algumas na outra extremidade formando uma longa cauda. Estatísticos desenvolveram uma ampla teoria sobre um tipo particular de distribuição simétrica chamada **distribuição normal** (Fig. 25.5). Neste modo de distribuição em forma de sino, a média e a classe modal estão situadas exatamente no centro. Em geral, as distribuições dos dados da amostra aproximam-se da forma de uma distribuição normal. Assim, podemos aplicar a ampla teoria sobre distribuições normais para analisar tais dados.

### A Variância e o Desvio Padrão

Os dados em uma distribuição de frequência podem ser dispersos ou podem estar agrupados. Para medir a distribuição dos dados em uma distribuição de frequência, usamos uma estatística chamada **variância**. Os dados que estão amplamente dispersos produzem um grande valor para a variância, enquanto os dados que estão muito agrupados produzem um pequeno valor. A variância da amostra, indicada por  $s^2$ , é calculada pela fórmula

$$s^2 = \Sigma (X_k - \bar{X})^2 / (n - 1)$$

Nesta fórmula,  $(X_k - \bar{X})^2$  é o quadrado da diferença entre a observação  $k^{\text{a}}$  e a média da amostra (geralmente chamada de *desvio quadrado da média*), e a letra grega  $\Sigma$  indica que todos os desvios quadrados estão somados. A soma dos desvios quadrados é obtida dividindo-se por  $n - 1$ . (Por motivos técnicos, o divisor é um a menos que o tamanho da amostra.)



Devemos notar duas características da variância. Primeira, ela mede a dispersão de dados ao redor da *média*. Quando nós calculamos a variância, tomamos a média para ser o valor central da distribuição e encontramos a diferença entre ela e cada uma das observações na amostra. Segunda, a variância é sempre positiva. Quando calculamos a variância, determinamos o quadrado da diferença entre cada observação e a média e então somamos os quadrados das diferenças. Como cada um dos quadrados das diferenças é positivo, a variância, calculada pela soma dos quadrados das diferenças, também é positiva.

Embora a variância tenha propriedades matemáticas desejáveis, ela é difícil de interpretar porque as unidades de medidas estão ao quadrado (por exemplo,  $s^2 = 2,88 \text{ dias}^2$ ). Consequentemente, uma outra estatística, chamada **desvio padrão**, em geral é usada para descrever a variabilidade de uma amostra. O desvio padrão ( $s$ ) é a raiz quadrada da variância da amostra

$$s = \sqrt{s^2}$$

Esta estatística é mais fácil de interpretar do que a variância porque é expressa nas mesmas unidades que as medidas originais.

As variâncias e desvios padrões das quatro populações de trigo são fornecidas na Figura 25.4. A população  $F_2$  tem a maior variância e o maior desvio padrão, sem dúvida porque está segregando alelos que controlam o tempo de maturação. Nas plantas de  $F_2$ , tanto diferenças genéticas quanto ambientais produzem a variabilidade observada. Nas outras populações, a maioria da variação observada ou toda ela se deve apenas a fatores ambientais. Cada uma das duas linhagens parentais é altamente endocruzada, e espera-se, portanto, que sejam homozigotas para a maioria de seus genes. As plantas de  $F_1$  são heterozigotas para os alelos que são diferentes nas duas linhagens parentais, mas todas elas têm o mesmo genótipo. Assim, nem nas linhagens parentais nem na  $F_1$  esperamos encontrar muita variação genética entre plantas. Em uma seção mais adiante, veremos como avaliar esta parte da variância em uma característica quantitativa que resulta de diferenças genéticas entre indivíduos em uma população.

Como mencionado anteriormente, a distribuição de uma característica quantitativa em geral parece uma distribuição normal. A forma e a posição de uma distribuição normal são totalmente especificadas por sua média e seu desvio padrão. Assim, se soubermos apenas a média e o desvio padrão de uma característica quantitativa e supusermos que a característica tem distribuição normal, podemos construir a forma aproximada da distribuição da característica. Nesta distribuição, 66% das medidas ficarão dentro de um desvio padrão da média, 95% ficarão dentro de dois desvios padrões da média e 99% ficarão dentro de três desvios padrões da média (Fig. 25.5).

### PONTOS IMPORTANTES

- A média ( $\bar{X} = (\sum X_k)/n$ ) e a classe modal apontam para o centro de uma distribuição de frequência.

- A variância ( $s^2 = \sum (X_k - \bar{X})^2 / (n - 1)$ ) e o desvio padrão  $s = \sqrt{s^2}$  são estatísticas que indicam a extensão na qual os dados estão distribuídos ao redor da média em uma distribuição de frequência.

## ANÁLISE DE CARACTERÍSTICAS QUANTITATIVAS

Os geneticistas quantitativos enfocam suas análises na variabilidade fenotípica medida pela variância.

Nesta seção, veremos como as estatísticas são usadas na análise genética de características quantitativas. O objetivo da análise é a divisão da variação observada na característica em componentes genético e ambiental e então o emprego do componente genético para fazer previsões sobre os fenótipos da prole de determinados cruzamentos.

### A Hipótese de Fatores Múltiplos

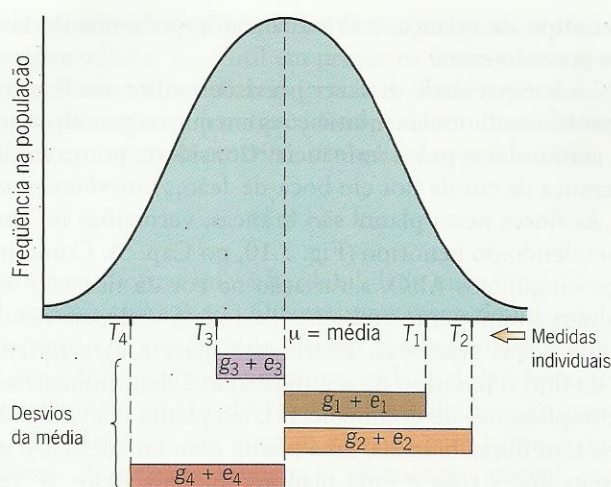
A idéia principal na genética quantitativa é que as características são controladas por muitos fatores diferentes no ambiente e no genótipo. Esta **hipótese de fatores múltiplos** emergiu na segunda década do século vinte pelas investigações experimentais de E. M. East, W. Johannsen, H. Nilsson-Ehle e outros. Entretanto, foi um teórico, R. A. Fisher, quem cristalizou a hipótese de fatores múltiplos em sua forma moderna. Fisher fez este trabalho durante a Primeira Guerra Mundial, enquanto ele estava ensinando em uma escola na Grã-Bretanha. Sua análise teórica foi publicada em 1918, o ano em que a guerra acabou.

Fisher formulou a hipótese de que um determinado valor de uma característica quantitativa,  $T$ , é o resultado da influência combinada de fatores genéticos e ambientais. Ele representou os efeitos destes fatores como desvios da média geral da população:

$$T = \mu + g + e$$

Nesta equação, a letra grega  $\mu$  representa a média da população,  $g$  representa o desvio da média que resulta de fatores genéticos e  $e$  representa o desvio da média que é decorrente de fatores ambientais. No esquema de Fisher, a posição de um determinado valor da característica,  $T$ , na população depende dos fatores genéticos e ambientais que a afetaram (Fig. 25.6). Alguns fatores produzem grandes valores de  $T$  e alguns produzem pequenos valores de  $T$ . Para cada indivíduo, estes fatores são diferentes. Além disso, Fisher destacou que vários fatores estão envolvidos. Ele formulou a hipótese de que muitos genes contribuem para uma característica quantitativa e acreditou que muitos aspectos do ambiente também fazem contribuições. Hoje em dia, dizemos





**Fig. 25.6** ■ Fenótipos quantitativos e os desvios de medidas individuais da média da população. Cada desvio individual supostamente consiste em um desvio devido a seu genótipo ( $g$ ) e um desvio devido a seu ambiente ( $e$ ).

que uma característica que é controlada por muitos genes é **poligênica**.\*

## Partição da Variância Fenotípica

Com estas idéias simples, Fisher foi capaz de desenvolver um procedimento para analisar a variabilidade de uma característica quantitativa em termos de fatores contribuintes genéticos e ambientais. Para medir a variabilidade da característica, ele enfocou a estatística que chamamos de *variância*. Especificamente, ele descobriu como dividir a variância geral da característica em dois componentes da variância, um que mede os efeitos das diferenças genéticas entre indivíduos e outro que mede os efeitos das diferenças ambientais. Assim, na análise de Fisher, a variância de uma característica quantitativa, representada por  $V_T$ , é igual à soma de uma *variância genética*, representada por  $V_g$ , e uma *variância ambiental*, representada por  $V_e$ :

$$V_T = V_g + V_e$$

Nesta equação de variância, a variância da característica quantitativa,  $V_T$ , geralmente é referida como *variância fenotípica total*.

Uma discussão sobre o método de Fisher de dividir a variância fenotípica total em seus componentes genético e ambiental está além do escopo deste livro. Entretanto, este método tem sido usado em muitos contextos diferentes e deu origem a uma técnica estatística geral chamada de *análise de variância*.

Para entender a idéia básica, vamos partir a variância do tempo de maturação na população  $F_2$  de trigo mostrada na

Figura 25.4. A variância fenotípica total desta população ( $V_T$ ) é  $14,26 \text{ dias}^2$ . Em termos da equação da variância de Fisher, este total pode ser representado como a soma de uma variância genética ( $V_g$ ) e uma variância ambiental ( $V_e$ ), ambas as quais devem ser estimadas usando-se outros dados. Para estimar a variância ambiental, podemos usar os dados das populações parental e  $F_1$ . As populações parentais são geneticamente uniformes porque são ambas endogâmicas. A população  $F_1$  também é geneticamente uniforme porque foi criada por cruzamento das duas populações endocruzadas. Espera-se que todas as plantas de  $F_1$  sejam identicamente heterozigotas para os genes que diferem nas populações parentais endocruzadas. Devido a esta uniformidade genética, a variabilidade que vemos em cada uma destas três populações deve refletir diferenças resultantes de efeitos ambientais. Para obter um valor representativo para  $V_e$ , podemos fazer a média das variâncias destes grupos:

$$\begin{aligned} V_e &= (V_A + V_B + V_{F1})/3 \\ &= (1,92 \text{ dia}^2 + 2,05 \text{ dias}^2 + 2,88 \text{ dias}^2)/3 \\ &= 2,28 \text{ dias}^2 \end{aligned}$$

Com esta estimativa da variância ambiental, podemos agora estimar  $V_g$  por subtração da variância total  $V_T$ :

$$\begin{aligned} V_g &= V_T - V_e \\ &= 14,26 \text{ dias}^2 - 2,28 \text{ dias}^2 \\ &= 11,98 \text{ dias}^2 \end{aligned}$$

Assim, a variância fenotípica total para o tempo de maturação na  $F_2$  da população de trigo foi dividida em dois componentes:

$$\begin{aligned} V_T &= V_g + V_e \\ 14,26 \text{ dias}^2 &= 11,95 \text{ dias}^2 + 2,28 \text{ dias}^2 \end{aligned}$$

A partir desta partição, vemos que a maior parte da variância no tempo de maturação da população  $F_2$  de trigo se deve a diferenças genéticas entre os indivíduos. Essa variabilidade genética surgiu da segregação e distribuição de genes quando as plantas de  $F_1$  foram reproduzidas. Tais plantas eram heterozigotas para genes que diferem nas populações parentais. Quando elas se reproduziram, a segregação e a distribuição produziram uma gama de genótipos – três genótipos distintos para cada gene heterozigoto. A variação que vemos na  $F_2$  resulta primariamente de diferenças fenotípicas entre esses genótipos.

## Herdabilidade em Sentido Amplo

Geralmente, é informativo calcular a proporção da variância fenotípica total que é oriunda de diferenças genéticas entre indivíduos em uma população. Esta proporção é chamada de **herdabilidade em sentido amplo**, representada por  $H^2$ . Em termos dos componentes da variância de Fisher,

$$\begin{aligned} H^2 &= V_g / V_T \\ &= V_g / (V_g + V_e) \end{aligned}$$

\*N.T.: Poligênico não é sinônimo de multifatorial.



O símbolo da herdabilidade em sentido amplo,  $H^2$ , é escrito com o expoente 2 para nos lembrar que esta estatística é calculada a partir de variâncias, que são quantidades ao quadrado.

Em virtude do modo como ela é calculada, a herdabilidade em sentido amplo deve situar-se entre 0 e 1. Se for próxima de 0, pouco da variabilidade observada na população é atribuível a diferenças genéticas entre indivíduos. Se for próxima de 1, a maior parte da variabilidade observada é atribuível a diferenças genéticas. A herdabilidade em sentido amplo resume, portanto, as contribuições relativas de fatores genéticos e ambientais para a variabilidade observada em uma população. Entretanto, é importante notar que esta estatística é específica da população. Para uma determinada característica, populações diferentes podem ter valores diferentes da herdabilidade em sentido amplo. Assim, a herdabilidade em sentido amplo de uma população não pode ser automaticamente suposta como representativa da herdabilidade em sentido amplo de outra população.

Na população  $F_2$  de trigo,  $H^2 = 11,98/14,26 = 0,84$ . Este resultado nos diz que nesta população 84% da variabilidade observada no tempo de maturação do trigo se deve a diferenças genéticas entre indivíduos. Entretanto, não nos diz o que são estas diferenças. A variância genética da qual a herdabilidade depende em sentido amplo inclui todos os fatores que fazem com que os genótipos tenham fenótipos diferentes: os efeitos de alelo individuais, as relações de dominância entre alelos e as interações epistáticas entre genes diferentes. No Cap. 4, vimos como tais fatores influenciam os fenótipos. Nas duas seções seguintes, veremos que, separando estes componentes de variabilidade genética e enfocando os componentes que envolvem os efeitos de alelos individuais, podemos prever os fenótipos da prole a partir dos fenótipos de seus genitores.

### Herdabilidade em Sentido Restrito

A capacidade de fazer previsões em genética quantitativa depende da quantidade de variação genética que decorre dos efeitos de alelos individuais. A variação genética que resulta dos efeitos de dominância e epistasia tem pouco poder preditivo.

Para ver como a dominância limita a capacidade de fazer previsões, considere os tipos sanguíneos ABO em humanos (Quadro 4.1, no Cap. 4). Esta característica é determinada estritamente pelo genótipo; a variação ambiental não tem efeito sobre o fenótipo. Entretanto, por causa da dominância, dois indivíduos com o mesmo fenótipo podem ter genótipos diferentes. Por exemplo, uma pessoa com grupo sanguíneo tipo A pode ser ou  $I^A I^A$  ou  $I^A i$ . Se duas pessoas com sangue tipo A produzem um filho, não podemos prever exatamente que fenótipo a criança terá. Ela poderá ser do tipo A ou do tipo O, dependendo dos genótipos dos genitores. Entretanto, sabemos que ela não terá tipo B ou tipo AB. Assim, embora possamos fazer algum tipo de previsão sobre

o fenótipo da criança, a dominância impede-nos de fazer uma previsão exata.

Nossa capacidade de fazer previsões sobre um fenótipo da prole é melhorada em situações em que os genótipos não são confundidos por dominância. Considere, por exemplo, a herança de cor da flor em boca-de-leão, *Antirrhinum majus*. As flores nesta planta são brancas, vermelhas ou rosa, dependendo do genótipo (Fig. 3.10, no Cap. 3). Como nos tipos sanguíneos ABO, a variação na cor da flor não tem nenhum componente ambiental, e toda a variância resulta de diferenças genéticas. Entretanto, para a característica cor da flor, o fenótipo de um indivíduo é determinado sem as complicações de dominância. Uma planta com dois alelos  $w$  tem flores brancas, uma planta com um alelo  $w$  e um  $W$  tem flores rosa e uma planta com dois alelos  $W$  tem flores vermelhas. Neste sistema, o fenótipo depende simplesmente do número de alelos  $W$  presentes. Cada alelo  $W$  intensifica a cor em uma quantidade fixa. Assim, podemos dizer que os alelos que contribuem para a determinação da cor contribuem para o fenótipo de modo estritamente aditivo. Essa aditividade melhora nossa capacidade de fazer previsões em cruzamentos entre plantas diferentes. Um cruzamento entre duas plantas vermelhas produz apenas prole vermelha; um cruzamento entre duas plantas brancas produz apenas prole branca; e um cruzamento entre plantas vermelhas e brancas produz apenas prole rosa. A única incerteza é em um cruzamento entre dois heterozigotos, e neste caso a incerteza é decorrente de segregação mendeliana, e não de dominância.

Os geneticistas quantitativos distinguem variância genética resultante de alelos que atuam aditivamente (como estes no exemplo de cor da flor que citamos) de variância genética que é oriunda de dominância. Estes componentes da variância são representados como:

$V_a$  = variância genética aditiva

$V_d$  = variância de dominância

Além disso, os geneticistas definem um terceiro componente da variância que mede a variação devida a interações epistáticas entre alelos de genes diferentes:

$V_i$  = variância epistática

Interações epistáticas, como a dominância, são de pouca ajuda na previsão de fenótipos.

Juntos, estes três componentes da variância constituem a variância genética total:

$$V_g = V_a + V_d + V_i$$

Se lembrarmos que  $V_T = V_g + V_e$ , podemos expressar a variância fenotípica total como a soma de quatro componentes:

$$V_T = V_a + V_d + V_i + V_e$$



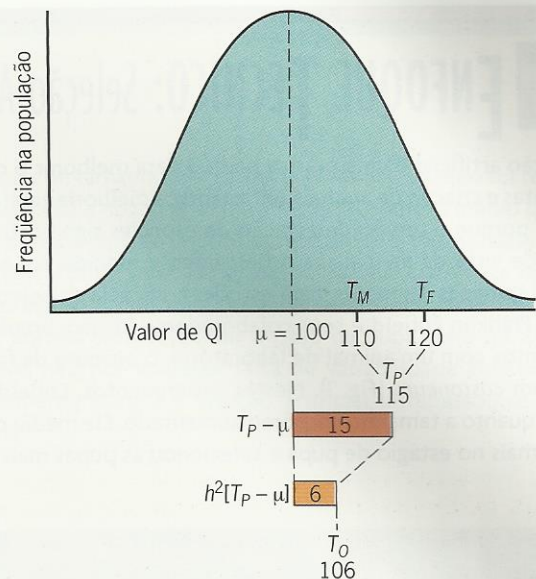
Destes quatro componentes da variância, apenas a variância genética aditiva,  $V_a$ , é útil em prever os fenótipos da prole a partir dos fenótipos de seus genitores. Esta variância, como fração da variância fenotípica total, é denominada **herdabilidade em sentido restrito**, simbolizada por  $b^2$ . Assim,

$$b^2 = V_a / V_T$$

Da mesma forma que a herdabilidade em sentido amplo,  $b^2$  situa-se entre 0 e 1. Quanto mais perto de um, maior a proporção da variância fenotípica total que é variância genética aditiva e maior nossa capacidade de prever o fenótipo da prole. O Quadro 25.1 fornece algumas estimativas da herdabilidade em sentido restrito para várias características. A estatura humana é altamente variável, mas o tamanho da prole em porcos não é. Assim, se soubermos os fenótipos parentais, seremos mais capazes de prever a altura de uma prole humana do que o tamanho dos filhotes de um porco.

### Previsão de Fenótipos

Para compreender melhor o significado da herdabilidade em sentido restrito, consideremos a situação diagramada na Figura 25.7. Michael (M) e Frances (F) fizeram um teste padrão de inteligência, e foi determinado seu coeficiente de inteligência (QI). O valor de Michael é 110 e o de Frances é 120. O valor médio de QI da população é 100. Michael e Frances tiveram um filho, Oswald (O), que foi dado para adoção quando nasceu, e os pais adotivos quiseram prever o QI de Oswald. Se o QI não tivesse componente genético, nossa melhor estimativa para o QI de Oswald seria 100, a média da população. Não temos meios para prever que tipo de ambiente domiciliar Oswald receberá e portanto não podemos prever que tipos de fatores não genéticos irão influenciar seu desenvolvimento mental. Nem podemos usar o QI de Michael e Frances para prever nada sobre o QI de Oswald, pois, por suposição, os genes que deram a ele não teriam nada a ver com o desenvolvimento mental. Entretanto, vários estudos indicaram que a variação nos valores de



**Fig. 25.7** ■ Previsão de um fenótipo da prole com base nos fenótipos de seus genitores e na herdabilidade em sentido restrito da característica. Apenas parte do desvio do valor médio parental ( $T_P$ ) da média da população é herdável. A magnitude desta parte é determinada pela herdabilidade em sentido restrito.

QI tem um componente genético. De fato, a herdabilidade em sentido restrito do QI foi estimada em cerca de 0,4, isto é, cerca de 40% da variação observada nos valores de QI se deve aos efeitos aditivos de alelos. Podemos usar esta estatística juntamente com os QI parentais para prever o valor de QI de Oswald?

Vamos simbolizar os QI de Oswald, Michael e Frances como  $T_O$ ,  $T_M$  e  $T_F$ , respectivamente, e representar a média da população como  $\mu$ . A melhor previsão do QI de Oswald é

$$T_O = \mu + b^2 [(T_M + T_F)/2 - \mu]$$

A expressão entre parênteses,  $(T_M + T_F)/2$ , é geralmente denominada *valor médio parental*. E a média dos fenótipos dos dois genitores. Se representarmos o valor médio parental com o símbolo  $T_P$ , a equação de previsão para o fenótipo de Oswald fica simplificada para

$$T_O = \mu + b^2 [T_P - \mu]$$

A expressão entre colchetes,  $[T_P - \mu]$ , é a diferença entre o valor médio parental e a média da população. O produto desta diferença e a herdabilidade em sentido restrito é o desvio previsto do fenótipo da prole da média da população. De fato, a herdabilidade no sentido restrito traduz a diferença entre o valor médio parental e a média da população em uma diferença “herdável” que podemos esperar ver na prole. Adicionando esta diferença herdável à média, podemos prever o fenótipo da prole.

#### QUADRO 25.1

##### Estimativas de Herdabilidade em Sentido Restrito ( $b^2$ ) para Características Quantitativas

| Característica                        | $b^2$ |
|---------------------------------------|-------|
| Estatura em seres humanos             | 0,65  |
| Produção de leite em gado             | 0,35  |
| Tamanho da prole em porcos            | 0,05  |
| Produção de ovos em galinhas          | 0,10  |
| Tamanho da cauda em camundongos       | 0,40  |
| Tamanho do corpo em <i>Drosophila</i> | 0,40  |

Fonte: D. S. Falconer. 1981. *Introduction to Quantitative Genetics*. 2ª ed., p. 51. Longman, London.





## ENFOQUE TÉCNICO: Seleção Artificial

A seleção artificial é uma prática padrão para melhorar o cultivo de plantas e criação de animais. Entretanto, a melhoria geralmente é lenta porque o tempo de geração de espécies significativas do ponto de vista da agricultura é tipicamente medido em anos, e não em semanas ou meses. Para estudar a eficácia da seleção artificial, Franklin Enfield e seus colaboradores fizeram amplos experimentos com um animal de laboratório, o besouro da farinha, *Tribolium castaneum* (Fig. 1). Nestes experimentos, Enfield selecionou quanto a tamanho corpóreo aumentado. Ele mediu o peso dos animais no estágio de pupa e selecionou as pupas mais pesa-



**Fig. 1** ■ Tamanhos contrastantes do besouro da farinha. O besouro menor é de um estoque padrão de um laboratório. O besouro maior mostra o resultado de mais de 120 gerações de seleção para tamanho corpóreo aumentado.

das para serem genitoras da geração seguinte. Este processo foi continuado por 125 gerações. No começo do experimento, o peso das pupas individualmente variava de 1.800 a 3.000  $\mu\text{g}$ , a média era de 2.400  $\mu\text{g}$  e a variância era de 40.000  $\mu\text{g}^2$ . Após 125 gerações de seleção, o peso médio de pupas tinha aumentado para 5.800  $\mu\text{g}$ , mais que o dobro da média da população inicial (Fig. 2a). Além disso, nenhum dos indivíduos na população selecionada era tão pequeno quanto os maiores indivíduos na população inicial original (Fig. 2b). Esta completa falta de superposição nas distribuições de frequência indica que a constituição genética da população foi radicalmente alterada.

Para obter este incrível resultado, Enfield usou uma seleção diferencial de 200  $\mu\text{g}$  em cada geração. Inicialmente, a herdabilidade em sentido restrito para o peso das pupas foi estimada em cerca de 0,3. Assim, a resposta prevista a seleção foi de  $0,3 \times 200 \mu\text{g} = 60 \mu\text{g}$  por geração (veja Fig. 25.8). Para as primeiras 40 gerações, isso foi aproximadamente o que Enfield observou. Entretanto, a resposta cumulativa durante essa época foi 2.000  $\mu\text{g}$ , um pouco menos que os 2.400  $\mu\text{g}$  esperados ( $60 \mu\text{g}/\text{geração} \times 40$  gerações). Esta discrepância foi devida a fatores que reduziram a eficiência da seleção, incluindo variáveis como infertilidade entre os indivíduos selecionados. Assim, embora a herdabilidade em sentido restrito seja um preditor razoavelmente bom da resposta a seleção durante algumas gerações, a longo prazo ela tende a superestimar essa resposta.

As últimas gerações do projeto de Enfield demonstraram marcadamente este ponto. Entre as gerações 40 e 125, a resposta cumulativa foi de 1.400  $\mu\text{g}$ , que, embora seja impressionante, é muito menor que a resposta esperada de 5.100  $\mu\text{g}$  ( $60 \mu\text{g}/\text{geração} \times 85$  gerações). Enfield verificou a herdabilidade em sentido restrito nestas últimas gerações para ver se parte da variância ge-

Vamos agora substituir as quantidades conhecidas para cada um dos termos na equação prevista:  $\mu = 100$ ,  $T_P = (110 + 120)/2 = 115$  e  $b^2 = 0,4$ . Assim, o valor previsto para  $T_O$  é

$$\begin{aligned} T_O &= 100 + (0,4)[115 - 100] \\ &= 106 \end{aligned}$$

Este resultado nos diz que o QI de Oswald é esperado situar-se entre o valor médio parental (115) e a média da população (100). De fato, ele está em um ponto a 40% da distância entre a média populacional e o valor médio parental. Esses 40% correspondem à herdabilidade em sentido restrito (0,4). Se a herdabilidade em sentido restrito do QI for maior do que 0,4, o valor previsto para o QI de Oswald seria mais próximo do valor médio parental. Para uma característica perfeitamente herdável,  $b^2 = 1$  e o valor previsto para o fenótipo da prole seria igual à mé-

dia dos dois fenótipos parentais. Assim, a herdabilidade em sentido restrito é uma estatística crítica. Ela nos diz a proximidade com que a prole se assemelha à média de seus genitores. Devemos destacar, entretanto, que o valor do QI que calculamos para Oswald é um valor preditivo, e não um conhecido com certeza. Se fôssemos procurar milhares de casais, cada um tendo um QI médio parental de valor 115, seria esperado que os QI de seus filhos formariam uma distribuição de frequência. O QI médio desta distribuição seria de 106. Entretanto, a maioria das crianças teria valores de QI mais altos ou mais baixos, alguns até mesmo mais altos que os QI de ambos os genitores e alguns até mesmo mais baixos que a média populacional de 100. A variabilidade nesta distribuição vem da segregação mendeliana dos alelos que influenciam o QI e de fatores no ambiente. Se, por exemplo, Oswald for criado em um lar com pouco ou nenhum estímulo intelectual,

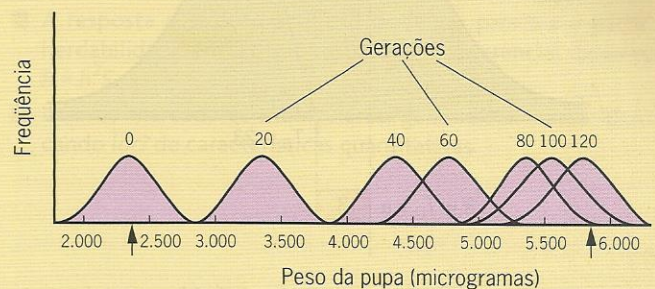
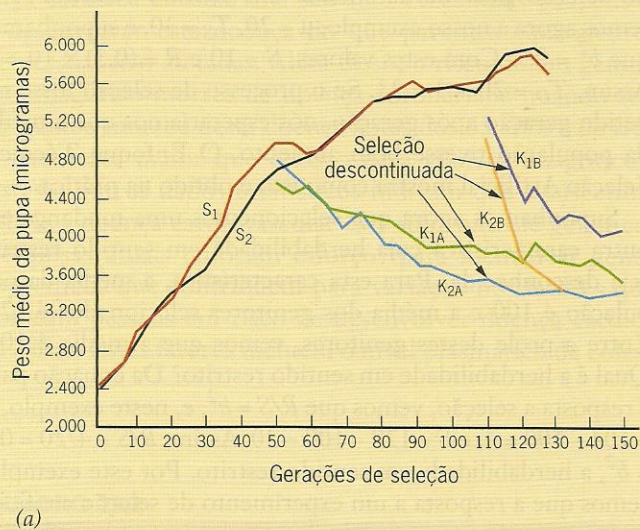




## ENFOQUE TÉCNICO: Seleção Artificial

nética aditiva tinha sido perdida durante o longo processo seletivo. Para sua surpresa, ele descobriu que  $h^2$  tinha sido apenas ligeiramente alterado, indicando que a população ainda reteve variabilidade genética selecionável. Uma análise mais detalhada demonstrou que, durante essas gerações, a eficiência da seleção foi gravemente reduzida por uma correlação negativa entre o tamanho e a capacidade reprodutiva. (Após um certo ponto, quanto maior o bezouro, menos ele tem sucesso reprodutivo.) Isto reduziu a seleção diferencial efetiva e dificultou selecionar aumentos

adicionais de tamanho. Na verdade, quando a seleção foi relaxada na geração 50 e novamente na geração 110, o peso médio da pupa começou a cair até um valor quase normal. Este resultado sugere que uma força de seleção natural foi substancialmente oposta à seleção artificial praticada no experimento principal. As tentativas de Enfield para aumentar o peso da pupa além de 5.800  $\mu\text{g}$  falharam, não por qualquer falta de variabilidade genética, mas simplesmente porque a seleção artificial ultrapassou os limites naturais.



**Fig. 2 ■** (a) Seleção artificial para tamanho aumentado no besouro da farinha, *Tribolium*. O tamanho foi medido pelo peso da pupa. As curvas  $S_1$  e  $S_2$  mostram a resposta a seleção em duas populações replicadas. As curvas  $K_{1A}$ ,  $K_{2A}$ ,  $K_{1B}$  e  $K_{2B}$  mostram o que aconteceu quando a seleção artificial foi descontinuada em subpopulações que foram estabelecidas a partir de populações selecionadas. (b) Distribuições de frequência de peso das pupas em populações de *Tribolium* selecionadas para aumento de tamanho. A forma das distribuições é apenas aproximada. As médias nas gerações 0 e 120 são indicadas por setas.

com uma dieta pobre e com outras condições desfavoráveis, seu QI pode ser consideravelmente menor que 106. Contrariamente, em um ambiente bem estimulante, o QI de Oswald pode ser muito maior que 106. Previmos que o QI de Oswald seria de 106; entretanto, devemos ter em mente que este número é uma previsão, e não um valor matematicamente estabelecido.

### Seleção Artificial

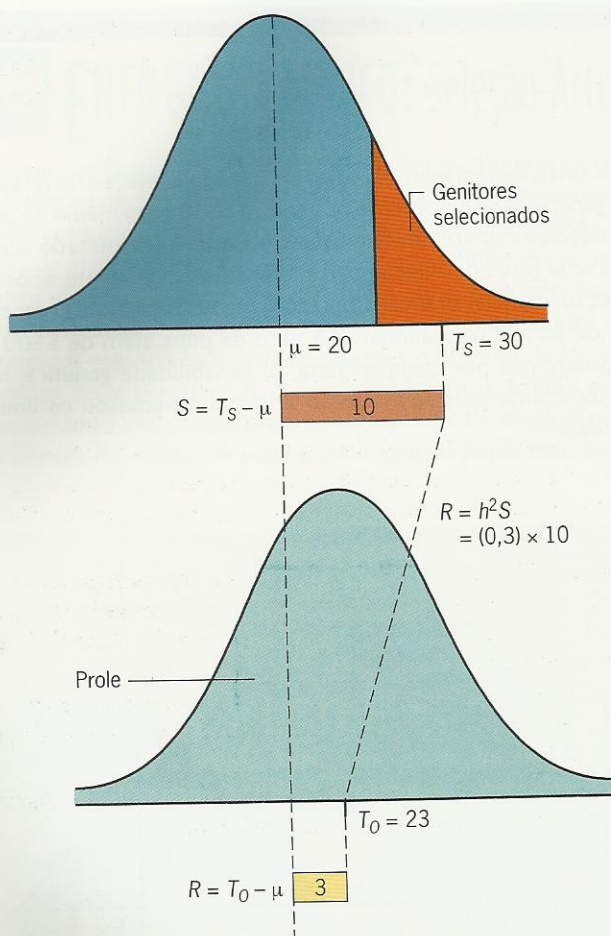
Além de prever um fenótipo da prole, a herdabilidade em sentido restrito tem outro emprego: prever o resultado de um programa de cruzamentos seletivos em uma população. As idéias estão resumidas na Figura 25.8, que mostra a distribuição de frequências de uma característica quantitativa entre genitores e sua prole. Na geração parental, o valor

médio da característica é 20. Para formar a geração seguinte, selecionamos os indivíduos na parte superior da distribuição para serem os genitores. Suponhamos que a média destes indivíduos selecionados seja de 30 unidades. Podemos prever o valor médio da característica na prole destes genitores selecionados? A resposta é sim, desde que saibamos a herdabilidade em sentido restrito da característica. A equação de previsão é

$$T_O = \mu + b^2[T_S - \mu]$$

onde  $T_O$  é a média da prole,  $\mu$  é a média da população geral,  $T_S$  é a média dos genitores selecionados e  $b^2$  é a herdabilidade em sentido restrito. Note que esta equação é a mesma equação prevista para o fenótipo de uma única prole, exceto que  $T_S$  foi substituído por  $T_P$ . Nós adaptamos a equação de previsão





**Fig. 25.8** ■ O processo de seleção artificial. O diferencial de seleção ( $S$ ) é a diferença entre a média dos genitores selecionados e a média da população. A resposta a seleção ( $R$ ) é a diferença entre a média da prole e a média da população geral que inclui seus genitores. A proporção  $R/S$  é igual à herdabilidade em sentido restrito.

de prole única para uma situação na qual muitos genitores (genitores *selecionados*) produzem um grupo inteiro de prole, que então forma a população na geração seguinte. Assim, a nova equação permite-nos prever como a média da população mudará mediante seleção dos indivíduos que serão genitores. Nós chamamos isto de **seleção artificial**. É prática comum em cruzamentos de animais e plantas e, em grande parte, é responsável por linhagens altamente produtivas de cultivos e espécies em fazendas que são usadas na agricultura de hoje.

Podemos ver mais claramente como a seleção muda a média de uma característica quantitativa em uma população reagrupando os termos na equação de seleção. Após subtrair  $\mu$  de ambos os lados da equação e introduzir colchetes ao redor do termo à esquerda, temos

$$[T_0 - \mu] = h^2[T_S - \mu]$$

O termo entre colchetes à direita,  $[T_S - \mu]$ , é chamado *diferencial de seleção*. Ele constitui a diferença entre a média

dos genitores selecionados e a média da população da qual eles foram selecionados. O diferencial de seleção mede a intensidade de seleção artificial. O termo entre colchetes à esquerda,  $[T_0 - \mu]$ , é chamado *resposta a seleção* e constitui a diferença entre a média da prole e a média de toda a população na geração anterior. Assim, a resposta a seleção mede o quanto a média da característica mudou em uma geração. Podemos colocar isto em termos ainda mais simples se representarmos a resposta a seleção por  $R$  e o diferencial de seleção por  $S$ . Então,

$$R = h^2 S$$

Assim, a resposta a seleção é o produto do diferencial de seleção pela herdabilidade em sentido restrito. Voltemos agora a nosso exemplo;  $\mu = 20$ ,  $T_S = 30$ , e suponhamos que  $h^2 = 0,3$ . Com estes valores,  $S = 10$  e  $R = (0,3) \times 10 = 3$ . Assim,  $T_0 = 20 + 3 = 23$ . Se o processo de seleção fosse repetido geração após geração, nós esperaríamos que a média da população tivesse bom aumento. O Enfoque Técnico: Seleção Artificial mostra como isto é obtido na prática.

Suponhamos agora que selecionemos uma mudança em outra característica cuja herdabilidade em sentido restrito seja desconhecida. Para esta característica, a média da população é 100 e a média dos genitores selecionados é 120. Entre a prole destes genitores, vemos que a média é 104. Qual é a herdabilidade em sentido restrito? Da equação para a resposta a seleção, vemos que  $R/S = h^2$ , e, neste exemplo,  $R = 104 - 100 = 4$ , e  $S = 120 - 100 = 20$ . Assim,  $R/S = 4/20 = 0,2 = h^2$ , a herdabilidade em sentido restrito. Por este exemplo, vemos que a resposta a um experimento de seleção artificial pode ser usada para estimar a herdabilidade em sentido restrito.

### Loci de Características Quantitativas

Análise estatística tem sido o principal eixo da genética quantitativa desde a publicação de Fisher em 1918. Com este tipo de análise, pesquisadores de genética quantitativa estudaram muitas características diferentes em muitos organismos diferentes e recentemente desenvolveram técnicas para identificar genes individuais que influenciam características complexas. Cada gene ocupa uma posição específica no conjunto de moléculas de DNA que constituem o genoma de um organismo. Tais moléculas de DNA são organizadas em estruturas distintas chamadas cromossomos. A posição de um gene em um cromossomo é chamada *locus* (plural, *loci*), e o *locus* de um gene que influencia uma característica quantitativa é chamado *locus de característica quantitativa*, abreviado como *locus QT* ou, mais simplesmente, *QTL*.

Técnicas moleculares modernas possibilitaram pesquisar os genomas quanto a *loci QT*. Consideramos algumas destas técnicas em capítulos anteriores. Os *loci QT* foram identificados e mapeados em cromossomos específicos em organismos modelos de laboratório tais como a mosca-das-frutas e o camundongo, em plantas de significado na agricultura



tais como milho e arroz, em animais tais como porcos e vacas e em nossa própria espécie. As características que foram estudadas incluem número de cerdas na mosca-das-frutas, obesidade no camundongo, produtividade no cultivo de arroz e milho, produção de leite no gado, gordura e taxa de crescimento em porcos e suscetibilidade a doenças tais como diabete, câncer, doenças cardiovasculares e esquizofrenia em seres humanos.

Começamos este capítulo com uma história sobre doença cardiovascular, que é a principal causa de morte entre pessoas nas sociedades pós-industriais. Sabemos há muito tempo que a suscetibilidade a esta doença é influenciada por fatores genéticos. Por exemplo, parentes que compartilham metade de seus genes com pessoas que têm doença cardíaca coronariana são sete vezes mais propensos a desenvolver esta doença do que parentes equivalentes de pessoas não afetadas. Além disso, o risco de um gêmeo monozigótico morrer de doença coronariana quando seu co-gêmeo morreu desta doença antes dos 65 anos é de três a sete vezes maior que o risco para gêmeos dizigóticos. Estes e outros dados estatísticos indicam que a suscetibilidade a doença cardiovascular está sob controle genético. Pesquisas atuais enfocam esforços para identificar genes específicos que contribuem para variação nos fatores que colocam pessoas em risco de desenvolver doença cardiovascular. Estes fatores incluem o nível plasmático de coles-

terol, obesidade, pressão arterial, níveis de lipoproteínas de alta e baixa densidade e nível de triglicerídeos. O Quadro 25.2 cita alguns dos *loci* QTL que foram identificados nesses esforços.

### PONTOS IMPORTANTES

- A variância fenotípica total pode ser dividida em componentes genético e ambiental:  $V_T = V_g + V_e$ .
- A variância fenotípica em uma população que é geneticamente uniforme avalia  $V_e$ .
- A herdabilidade em sentido amplo é a proporção da variância fenotípica total que é variância genética:  $H^2 = V_g / V_T$ .
- A variância genética pode ser subdividida em variâncias genéticas aditiva, de dominância e epistática:  $V_g = V_a + V_d + V_i$ .
- A herdabilidade em sentido restrito é a proporção da variância fenotípica total que é devida aos efeitos aditivos de alelos:  $h^2 = V_a / V_T$ .
- A herdabilidade em sentido restrito é usada para prever os fenótipos da prole ( $T_O$ ) considerando o fenótipo médio dos genitores ( $T_P$ ) e o fenótipo médio na população ( $\mu$ ) da qual vieram os genitores:  $T_O = \mu + h^2 (T_P - \mu)$ .
- A resposta a seleção artificial pode ser prevista a partir da herdabilidade em sentido restrito e do diferencial de seleção:  $R = h^2 S$ .
- Usando marcadores moleculares, geneticistas estão identificando *loci* de características quantitativas.

## QUADRO 25.2

**Loci de Características Quantitativas que Contribuem para Variação em Fatores de Risco de Doença Cardiovascular**

| Locus  | Produto Gênico                                   | Cromossomo | Fator de Risco                              |
|--------|--------------------------------------------------|------------|---------------------------------------------|
| AGT    | Angiotensina                                     | 1          | Pressão arterial                            |
| APOA-1 | Apolipoproteína A1                               | 11         | Colesterol HDL <sup>a</sup>                 |
| APOA-2 | Apolipoproteína A2                               | 1          | Colesterol HDL                              |
| APOA-4 | Apolipoproteína A4                               | 11         | Colesterol HDL, triglicerídeos              |
| APOB   | Apolipoproteína B                                | 2          | Colesterol LDL <sup>b</sup>                 |
| APOC-3 | Apolipoproteína C3                               | 11         | Triglicerídeos                              |
| APOE   | Apolipoproteína E                                | 19         | Colesterol LDL, triglicerídeos              |
| CETP   | Proteína de transferência de éster de colesterol | 16         | Colesterol HDL                              |
| DCP    | Dipeptidil-carboxipeptidase                      | 17         | Colesterol HDL, pressão arterial            |
| FGA/B  | Fibrinogênio A e B                               | 4          | Fibrinogênio                                |
| HRG    | Glicoproteína rica em histidina                  | 3          | Glicoproteína rica em histidina             |
| LDLR   | Receptor de lipoproteína de baixa densidade      | 19         | Colesterol LDL                              |
| LPA    | Lipoproteína (a)                                 | 6          | Colesterol HDL, triglicerídeos              |
| LPL    | Lipoproteína-lipase                              | 8          | Triglicerídeos                              |
| PLAT   | Ativador de plasminogênio tipo tissular          | 8          | Nível de ativador de plasminogênio tissular |
| PLANHI | Inibidor-1 de ativador de plasminogênio          | 7          | Nível de PAI-1                              |

Fonte: G. P. Vogler et al. 1997. Genetics and behavioral medicine: risk factors for cardiovascular disease. *Behavioral Medicine* 22:141-149.

<sup>a</sup>Lipoproteína de alta densidade.

<sup>b</sup>Lipoproteína de baixa densidade.

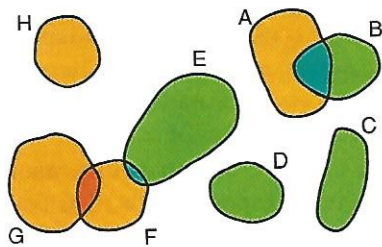


# Apêndice A

## Regras da Probabilidade

A teoria da probabilidade indica a frequência de eventos; por exemplo, a chance de obter o resultado cara ao jogar uma moeda, retirar um ás do baralho ou obter um homozigoto dominante ao cruzar dois heterozigotos. Em cada caso, o evento é o resultado de um processo – jogar uma moeda, retirar uma carta, gerar um filho. Para determinar a probabilidade de um evento específico, é preciso levar em conta todos os possíveis resultados do processo. O conjunto de todos os eventos é denominado *espaço amostral*. Ao jogar uma moeda, o espaço amostral contém dois eventos, cara e coroa; ao retirar uma carta, contém 52, um para cada carta; e a prole de heterozigotos tem três, GG, Gg e gg. A probabilidade de um evento é a frequência desse evento no espaço amostral. Por exemplo, as probabilidades associadas a cada descendente de um cruzamento entre dois heterozigotos são  $1/4$  (GG),  $1/2$  (Gg) e  $1/4$  (gg).

Dois tipos de questão surgem com frequência em problemas de probabilidade: (1) Qual é a probabilidade de que dois eventos, A e B, ocorram juntos? (2) Qual é a probabilidade de ocorrência de pelo menos um dos dois eventos, A ou B? A primeira questão especifica a ocorrência conjunta de dois eventos – é preciso que A e B ocorram juntos para satisfazer essa questão. A segunda questão é menos rigorosa – se ocorrer A ou B, a condição é satisfeita. Um diagrama simples ajuda a explicar os diferentes significados dessas duas questões.



As figuras no diagrama representam eventos no espaço amostral, e os tamanhos das figuras refletem suas frequências relativas. As superposições entre as figuras indicam a ocorrência conjunta de dois eventos. Se não houver superposição, eles nunca ocorrem juntos. A primeira questão verifica a probabilidade de ocorrência de

A e B; essa probabilidade é representada pelo tamanho da superposição dos dois eventos. A segunda questão verifica a probabilidade de ocorrência de A ou B; essa probabilidade é representada pela combinação das figuras dos dois eventos, inclusive, é claro, a superposição delas.

**A regra da multiplicação: se os eventos A e B forem independentes, a probabilidade de que ocorram juntos, denominada  $P(A \text{ e } B)$ , será  $P(A) \times P(B)$ .**

Aqui  $P(A)$  e  $P(B)$  são as probabilidades dos eventos individuais. Observe que independente não significa ausência de superposição no espaço amostral. Na verdade, os eventos sem superposição (disjuntos) não são independentes, já que se um ocorre, o outro não pode ocorrer. Em teoria de probabilidade, independente significa que um evento não fornece informações sobre o outro. Por exemplo, se uma carta retirada de um baralho for um ás, não temos indicação sobre o naipe da carta. Assim, a retirada de um ás de copas representa a ocorrência conjunta de dois eventos independentes – a carta é um ás (A) e é de copas (C). De acordo com a regra da multiplicação,  $P(A \text{ e } C) = P(A) \times P(C)$ , e como  $P(A) = 4/52$  e  $P(C) = 1/4$ ,  $P(A \text{ e } C) = (4/52) \times (1/4) = 1/52$ .

**A regra da adição: se os eventos A e B forem independentes, a probabilidade de que pelo menos um deles ocorra, denominada  $P(A \text{ ou } B)$ , será  $P(A) + P(B) - [P(A) \times P(B)]$ .**

Aqui o termo  $P(A) \times P(B)$ , que é a probabilidade de que A e B ocorram juntos, é subtraído da soma das probabilidades  $P(A) + P(B)$ , porque a soma inclui esse termo duas vezes. Como exemplo, suponha que queiramos saber a probabilidade de que uma carta retirada de um baralho seja um ás ou de copas. De acordo com a regra da adição,  $P(A \text{ ou } C) = P(A) + P(C) - [P(A) \times P(C)] = 4/52 + 1/4 - [(4/52) \times (1/4)] = 16/52$ .

Se não houver superposição dos dois eventos no espaço amostral, a expressão da regra da adição pode ser simplificada:  $P(A \text{ ou } B) = P(A) + P(B)$ . Por exemplo, suponha que queiramos saber a probabilidade de que uma carta retirada de um baralho seja um ás ou um rei (K). Não há superposição desses dois eventos no espaço amostral, isto é, eles são mutuamente exclusivos. Assim,  $P(A \text{ ou } K) = P(A) + P(K) = (4/52) + (4/52) = 8/52$ .



# Apêndice B

## Probabilidades Binomiais

Às vezes, a prole dos cruzamentos é segregada em duas classes distintas – por exemplo, macho ou fêmea, saudável ou doente, normal ou mutante, fenótipo dominante ou fenótipo recessivo. Para generalizar, podemos designar esses dois tipos de prole de P e Q e observar que, para qualquer indivíduo da prole, a probabilidade de ser P é  $p$  e a probabilidade de ser Q é  $q$ . Como só existem duas classes,  $q = 1 - p$ . Suponha que o número total de indivíduos da prole seja  $n$  e que a produção de cada um seja independente. Podemos calcular a **probabilidade binomial** de que exatamente  $x$  da prole pertença a uma classe e  $y$  à outra:

Probabilidade de  $x$  na classe P e de  $y$  na

$$\text{classe Q} = \left[ \frac{n!}{x! y!} \right] p^x q^y$$

O termo entre colchetes contém três funções fatoriais ( $n!$ ,  $x!$  e  $y!$ ) e cada uma delas é calculada como uma série decrescente de produtos. Por exemplo,  $n! = n(n-1)(n-2)(n-3)\dots(3)(2)(1)$ . O valor de  $0!$ , se necessário, é definido como um. Na fórmula, o termo entre colchetes, geralmente denominado **coeficiente binomial**, conta as diferentes formas, ou ordens, nas quais os  $n$  indivíduos da prole podem ser segregados de modo que  $x$  esteja na classe P e  $y$  esteja na classe Q. O outro termo,  $p^x q^y$ , indica a probabilidade de obter determinada forma ou ordem. Como todas as ordens são igualmente prováveis, a multiplicação desse termo pelo termo entre colchetes indica a probabilidade de obter prole  $x$  na classe P e  $y$  na classe Q, qualquer que seja a ordem de ocorrência.

Se, para valores fixos de  $n$ ,  $p$  e  $q$ , variarmos sistematicamente  $x$  e  $y$ , podemos calcular todo um conjunto de probabilidades. Esse conjunto constitui uma *distribuição binomial de probabilidade*. Com a distribuição, podemos responder a perguntas como “Qual é a probabilidade de que  $x$  ultrapasse determinado valor?” ou “Qual é a probabilidade de que  $x$  esteja entre dois valores específicos?” Por exemplo, consideremos uma família com seis filhos. Qual é a probabilidade de que pelo menos quatro sejam meninas? Para responder a essa pergunta, observemos que a probabilidade de que qualquer criança seja uma menina ( $p$ ) é  $1/2$  e a probabilidade de que seja um menino ( $q$ ) também é  $1/2$ . Portanto, a probabilidade de que exatamente quatro crianças de uma família sejam meninas (e dois sejam meninos) é  $[(6!)/(4! 2!)](1/2)^4 (1/2)^2 = 15/64$ , que é um dos termos na distribuição binomial.

Entretanto, a probabilidade de que pelo menos quatro sejam meninas (e no máximo dois sejam meninos) é a soma de três termos dessa distribuição:

| Evento                | Fórmula binomial                          | Probabilidade |
|-----------------------|-------------------------------------------|---------------|
| 4 meninas e 2 meninos | $[(6!)/(4! 2!)] \times (1/2)^4 (1/2)^2 =$ | 15/64         |
| 5 meninas e 1 menino  | $[(6!)/(5! 1!)] \times (1/2)^5 (1/2)^1 =$ | 6/64          |
| 6 meninas e 0 menino  | $[(6!)/(6! 0!)] \times (1/2)^6 (1/2)^0 =$ | 1/64          |

Portanto, a resposta é  $(15/64) + (6/64) + (1/64) = 22/64$ .

A distribuição binomial também oferece respostas a outros tipos de questões. Por exemplo, qual é a probabilidade de que pelo menos uma, mas no máximo quatro crianças sejam meninas? A resposta é a soma dos quatro termos:

| Evento                | Fórmula binomial                          | Probabilidade |
|-----------------------|-------------------------------------------|---------------|
| 1 menina e 5 meninos  | $[(6!)/(1! 5!)] \times (1/2)^1 (1/2)^5 =$ | 6/64          |
| 2 meninas e 4 meninos | $[(6!)/(2! 4!)] \times (1/2)^2 (1/2)^4 =$ | 15/64         |
| 3 meninas e 3 meninos | $[(6!)/(3! 3!)] \times (1/2)^3 (1/2)^3 =$ | 20/64         |
| 4 meninas e 2 meninos | $[(6!)/(4! 2!)] \times (1/2)^4 (1/2)^2 =$ | 15/64         |

Depois de somar, encontramos a resposta,  $56/64$ .

Consideremos o exemplo apresentado no Capítulo 3. Um homem e uma mulher, ambos heterozigotos para o alelo mutante recessivo causador da fibrose cística, pretendem ter quatro filhos. Qual é a chance de que um deles tenha fibrose cística e os outros três, não? Já vimos por enumeração que a resposta é  $108/256$  (ver Figura 3.14). No entanto, essa resposta também poderia ser obtida pela fórmula binomial. A probabilidade de que determinada criança seja afetada é  $p = 1/4$ , e a probabilidade de que não seja afetada é  $q = 3/4$ . O número total de crianças é  $n = 4$ , o número de crianças afetadas é  $x = 1$ , e o número de crianças não afetadas é  $y = 3$ . Reunindo todos esses dados, podemos calcular que a probabilidade de que exatamente um dos quatro filhos do casal tenha fibrose cística é

$$[4!/1! 3!] (1/4)^1 (3/4)^3 = 4 \times (1/4) \times (27/64) = 108/256$$



# Apêndice C

## Hibridização *in Situ*

Em 1969, Mary Lou Pardue e Joseph Gall desenvolveram um procedimento que possibilitava a hibridização de filamentos simples de DNA radioativos com filamentos complementares de DNA em cromossomos sobre lâminas de vidro. Com esse procedimento, denominado **hibridização *in situ***, Pardue e Gall puderam determinar as localizações cromossômicas de sequências repetitivas de DNA. (O termo latino *in situ* significa “em seu lugar original”; hibridização é a formação de moléculas duplas “híbridas” pelo pareamento de bases de filamentos complementares ou parcialmente complementares de DNA ou RNA.) A hibridização *in situ* clássica usava a dispersão de cromossomos mitóticos sobre lâmina de vidro (ver Figura 6.1), desnaturação de DNA nos cromossomos por exposição a álcali (NaOH 0,07 N) por alguns minutos, enxágue com tampão para remover a solução alcalina, incubação da lâmina em solução de hibridização contendo cópias radioativas da sequência nucleotídica de interesse, lavagem dos filamentos radioativos não hibridizados com sequências complementares nos cromossomos, exposição da lâmina a uma emulsão fotográfica sensível à radioatividade de baixa energia, revelação da autorradiografia e superposição da autorradiografia a uma fotografia dos cromossomos (Figura 1A).

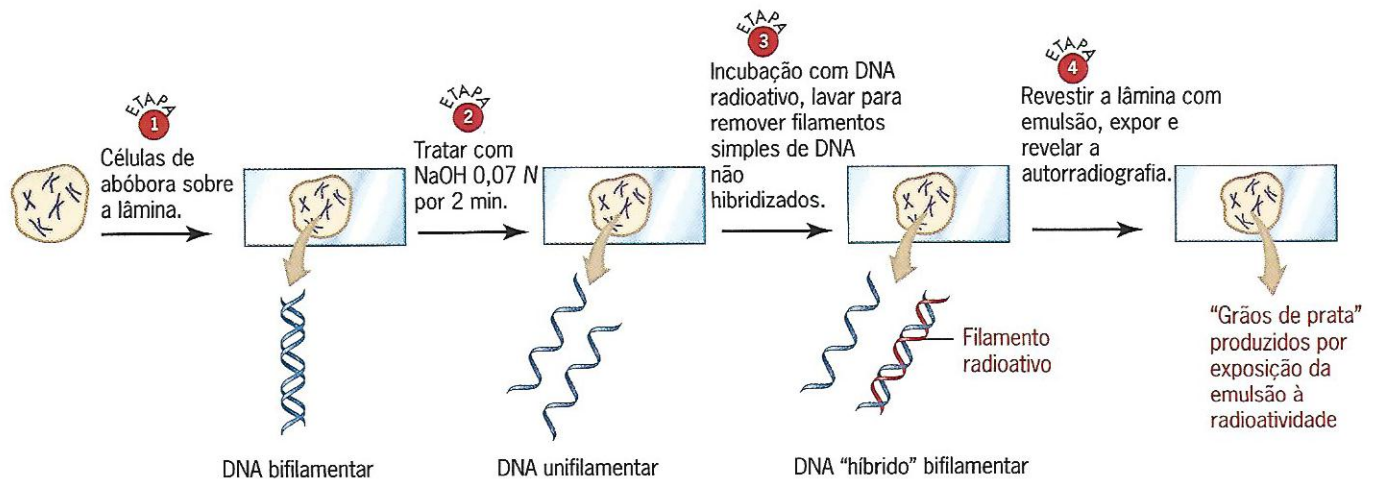
Um dos primeiros experimentos de hibridização *in situ* de Pardue e Gall mostrou que a sequência de DNA satélite do camundongo está localizada em regiões heterocromáticas que flanqueiam os centrômeros dos cromossomos de camundongo. O genoma de camundongo contém cerca de  $10^6$  cópias dessa sequência de DNA satélite, que tem cerca de 400 pares de nucleotídeos e consti-

tui cerca de 10% do genoma do camundongo. Em seguida, fizeram-se estudos semelhantes com os DNA satélites de várias outras espécies, e essas sequências de DNA repetitivo geralmente estão localizadas na heterocromatina centromérica ou adjacente aos telômeros.

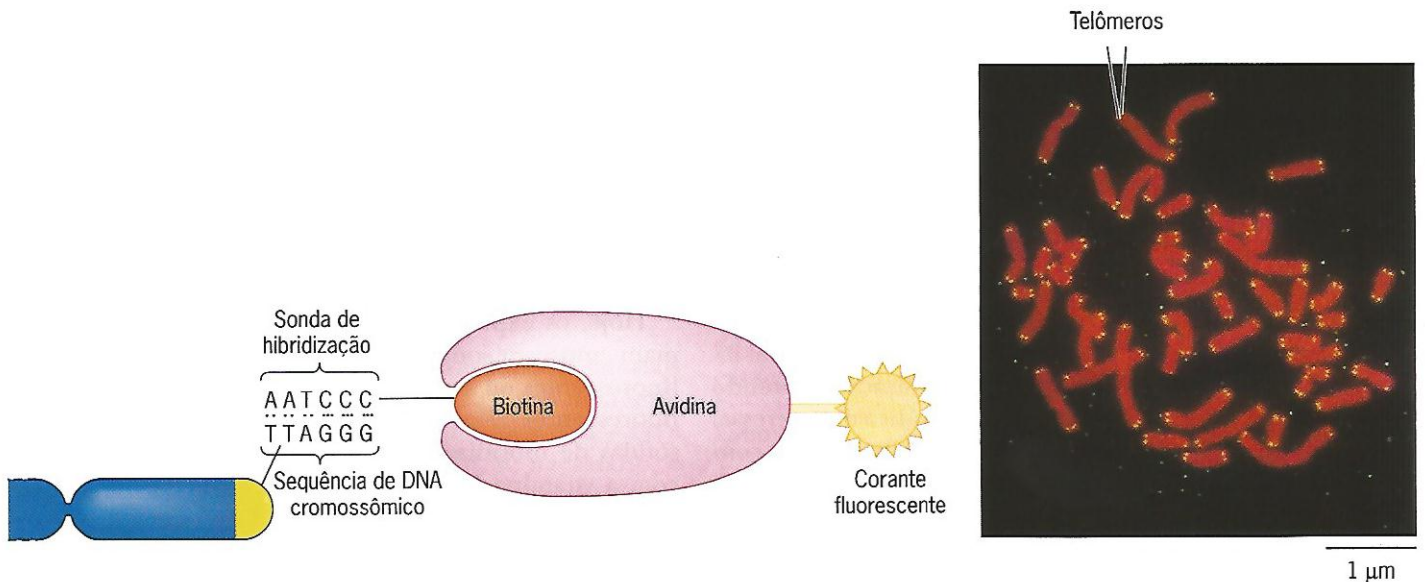
Uma sequência de DNA repetitivo só pode ser identificada como DNA satélite se a composição de bases for suficientemente diferente da composição do DNA da banda principal para produzir uma banda distinta durante a centrifugação por gradiente de densidade. Portanto, a centrifugação não pode ser usada para identificar todas as sequências de DNA repetitivo. As sequências de DNA satélite geralmente não são expressas; ou seja, não codificam RNA nem produtos proteicos.

Hoje, os experimentos de hibridização *in situ* costumam empregar sondas de hibridização ligadas a corantes fluorescentes ou a anticorpos marcados com substâncias fluorescentes (Figura 1B e 1C). Em um protocolo, as sondas de hibridização de DNA ou RNA são marcadas com a vitamina biotina, que se liga com alta afinidade à proteína do ovo avidina (Figura 1B). Com o uso da avidina unida por ligação covalente a um corante fluorescente, é possível detectar a localização cromossômica da sonda hibridizada pela fluorescência do corante. Esse procedimento, denominado **FISH (Hibridização *In Situ* Fluorescente)**, foi usado para demonstrar a presença da sequência repetitiva TTAGGG nos telômeros dos cromossomos humanos (Figura 1C). O FISH é muito sensível e pode ser usado para detectar as localizações de sequências de cópia única em cromossomos humanos em mitose e em intérfase.





A. Etapas da hibridização *in situ*.



B. Visualização de telômeros humanos usando corantes fluorescentes e hibridização *in situ*.

C. Telômeros humanos visualizados com o auxílio de sondas fluorescentes e hibridização *in situ*.

■ **FIGURA 1** Localização de seqüências repetidas de DNA em cromossomos por hibridização *in situ* realizada com sondas radioativas (A) ou sondas fluorescentes (B e C). O procedimento de hibridização *in situ* desenvolvido por Pardue e Gall é mostrado em (A). O uso de corantes fluorescentes para localizar a seqüência repetida TTAGGG nos telômeros de cromossomos humanos é ilustrado em (B), e uma fotomicrografia que mostra sua localização telomérica é mostrada em (C).



# Apêndice D

## Evidências de um RNA Mensageiro Instável

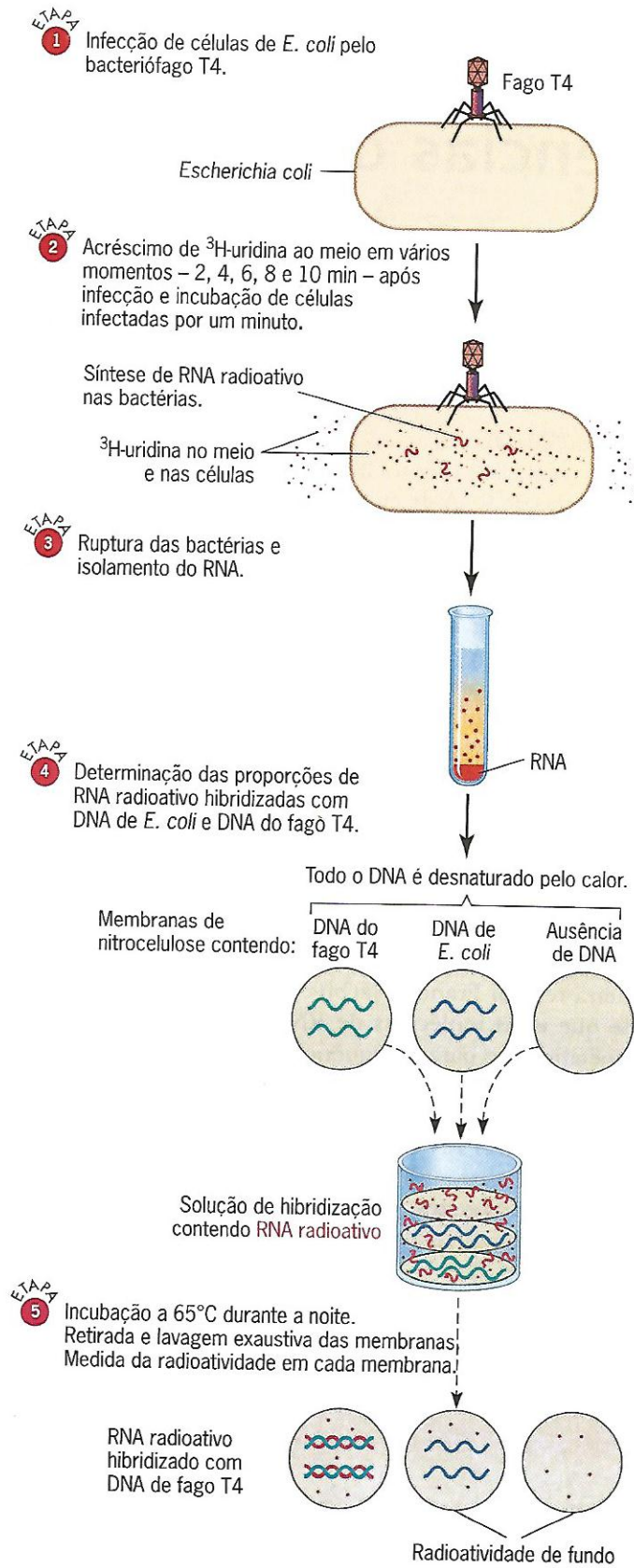
A primeira evidência da existência de um RNA intermediário na síntese de proteínas surgiu nos estudos de Elliot Volkin e Lawrence Astrachan em bactérias infectadas por vírus bacterianos. Os resultados, publicados em 1956, sugeriram que a síntese de proteínas virais em bactérias incluía a participação de moléculas de RNA instáveis especificadas por DNA viral. Volkin e Astrachan observaram uma explosão de síntese de RNA após infecção de células de *E. coli* pelo bacteriófago T2. Marcando o RNA com o isótopo radioativo  $^{32}\text{P}$ , eles mostraram que as moléculas de RNA recém-sintetizadas eram instáveis, com meia-vida de apenas alguns minutos. Além disso, mostraram que a composição nucleotídica dos RNA instáveis era semelhante à composição do DNA de T2 e diferente da do DNA de *E. coli*. Os resultados logo foram ampliados por estudos em outros laboratórios.

Em 1961, Sol Spiegelman e colaboradores relataram que os RNA instáveis sintetizados em células infectadas por fago T4 poderiam formar dúplex de RNA-DNA com DNA de T4 desnaturado, mas não com DNA de *E. coli* desnaturado. Eles marcaram bactérias com pulsos de  $^3\text{H}$ -uridina em vários momentos depois da infecção com o fago T4, isolaram todo o RNA dessas células e verificaram se as moléculas de RNA radioativo se hibridizaram com DNA de *E. coli* ou DNA de fago T4. O experimento é apresentado em diagrama na Figura 1.

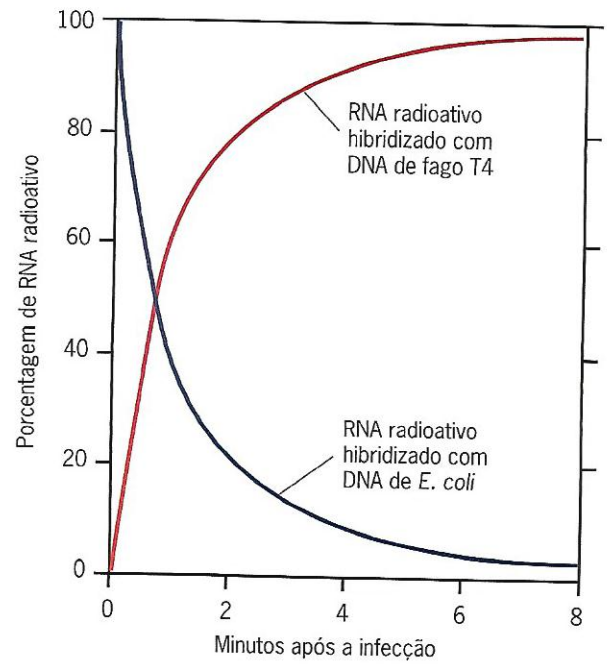
Os resultados (Figura 2) mostraram que a maioria das moléculas de RNA de curta duração sintetizadas após infecção era complementar a filamentos simples de DNA de fago T4 e não complementar a filamentos simples de DNA de *E. coli*. Esse achado indicou que elas eram produzidas a partir de moldes de DNA do fago T4, não de moldes de DNA de *E. coli*.

No mesmo ano em que Spiegelman e colegas publicaram seus resultados, Sydney Brenner, François Jacob e Matthew Meselson demonstraram que as proteínas do fago T4 eram sintetizadas em ribossomos de *E. coli*. Assim, as sequências de aminoácidos de proteínas T4 não eram controladas por componentes dos ribossomos. Em vez disso, os ribossomos eram as bancadas de trabalho nas quais ocorria a síntese proteica, mas não garantiam as especificações para proteínas individuais. Esses resultados reforçaram a ideia, proposta formalmente pela primeira vez por François Jacob e Jacques Monod em 1961, de que essas moléculas de RNA instáveis carregavam as especificações para as sequências de aminoácidos de produtos gênicos individuais dos genes para os ribossomos. As pesquisas subsequentes estabeleceram com firmeza o papel desses RNA instáveis, agora denominados RNA mensageiros ou mRNA, na transferência de informações genéticas dos genes para os sítios de síntese proteica no citoplasma.





■ FIGURA 1 Experimento de Spiegelman.



■ FIGURA 2 Rápida mudança da transcrição de genes de *E. coli* para genes de fago T4 em bactérias infectadas por T4.



# Apêndice E

## Taxas Evolutivas

As sequências de nucleotídeos e aminoácidos são os dados fundamentais para estudo da evolução molecular. Uma vez alinhadas as sequências homólogas de diferentes organismos, podemos determinar quantas posições nas moléculas são iguais ou diferentes; então, com a ajuda de dados de fósseis sobre a história dos organismos, podemos estimar a taxa de evolução molecular.

O caso mais simples é a comparação das sequências de aminoácidos de dois polipeptídios homólogos. Considere, por exemplo, os dois polipeptídios mostrados na Figura 1. Em três das quatro posições nesses dois polipeptídios, os aminoácidos são idênticos; na posição remanescente, eles são diferentes – glicina em um polipeptídio e serina no outro. Essa diferença de apenas um aminoácido indica que houve pelo menos uma substituição de aminoácido durante a evolução dos dois polipeptídios. O aminoácido ancestral pode ter sido a serina, caso em que a glicina em um polipeptídio representa um evento de substituição, ou o aminoácido ancestral pode ter sido a glicina, caso em que a serina no outro polipeptídio representa um evento de substituição.

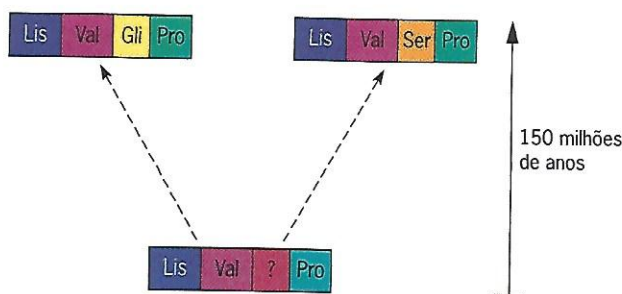
Entretanto, a história desses polipeptídios pode ser mais complicada. O aminoácido ancestral na posição variável pode ter sido outro que não a serina ou a glicina, por exemplo, arginina. Nesse caso, os dois polipeptídios descendentes sofreram substituições de aminoácidos durante sua evolução. Assim, o número mínimo de substituições de aminoácidos seria dois. Dizemos “mínimo” porque podem ter ocorrido múltiplas substituições na posição variável de qualquer um dos polipeptídios descendentes durante sua evolução. Desse modo, concentrando-se nas diferenças de aminoácidos em posições correspondentes em polipeptídios homólogos, não é possível contar o número real de substituições de aminoá-

cidos ocorridas. Tudo que podemos dizer é que houve, *no mínimo*, uma substituição desse tipo. Essa incerteza atrapalha a estimativa da taxa de evolução molecular, que, afinal, é o número total de substituições de aminoácidos ocorridas dividido pelo tempo total de evolução dos polipeptídios.

Para contornar esse problema, voltamos a atenção – paradoxalmente – para os aminoácidos que são iguais nos dois polipeptídios. Esses aminoácidos provavelmente não se modificaram em nenhum dos polipeptídios desde que as duas linhagens divergiram de um ancestral comum. Assim, oferecem informações sobre a probabilidade de que *não* ocorra uma substituição de aminoácido durante a evolução. Se pudermos estimar essa probabilidade, podemos contornar a situação e estimar a probabilidade de que ocorra uma substituição e, a partir dela, obter a taxa evolutiva.

Suponha que  $S$  seja a proporção de aminoácidos *iguais* nos dois polipeptídios – em nosso exemplo,  $S = 0,75$  – e suponha que  $v$  seja a probabilidade de que ocorra uma substituição de aminoácido em um sítio de um dos polipeptídios durante 1 ano do tempo evolutivo – ou seja,  $v$  é a taxa anual de substituição de aminoácido por sítio nesses polipeptídios. Ao definirmos  $v$  dessa forma,  $1 - v$  é a probabilidade de que não haja substituição de aminoácido em um sítio em qualquer ano do tempo evolutivo.

A partir do registro fóssil podemos determinar quando as duas linhagens que têm esses polipeptídios divergiram de um ancestral comum. Nos polipeptídios da Figura 1, essa divergência ocorreu há 150 milhões de anos. De modo geral, se o tempo desde a divergência do ancestral comum é de  $T$  anos, então o tempo evolutivo total das duas linhagens é  $T + T = 2T$  anos. Essa soma representa o número total de oportunidades anuais de que haja uma substituição de aminoácido em determinado sítio nos polipeptídios em evolução. Também representa o número total de oportunidades anuais de que *não* haja substituição nesse sítio. Assim, no fim do processo evolutivo, a probabilidade de que não tenha ocorrido uma substituição de aminoácido em determinado sítio em nenhum dos polipeptídios é o produto de todas as chances independentes individuais de que isso não ocorra, o que é igual a  $(1 - v)^{2T}$ . Em outras palavras, a probabilidade de que aminoácidos correspondentes nos dois polipeptídios tenham permanecido os mesmos durante o processo evolutivo é a probabilidade de que nenhum deles tenha se



■ FIGURA 1 Comparação de dois polipeptídios homólogos que evoluíram de maneira independente por 150 milhões de anos.



modificado em nenhum ano, que é  $(1 - v)^{2T}$ . Podemos estimar essa probabilidade pela proporção de aminoácidos que são iguais atualmente nos dois polipeptídios – isto é, por  $S$ . Assim,

$$S = (1 - v)^{2T}$$

Para encontrar  $v$ , a taxa anual de substituição de aminoácido por sítio, calculamos o logaritmo natural de ambos os lados da equação.

$$\begin{aligned}\ln S &= \ln(1 - v)^{2T} \\ \ln S &= 2T \ln(1 - v)\end{aligned}$$

Como  $v$  é um número muito pequeno – na verdade, bem próximo de zero –  $\ln(1 - v)$  é quase igual a  $-v$  (a curva logarítmica é quase linear quando o argumento da função logarítmica é próximo de 1). Assim,

$$\ln S = -2Tv$$

o que significa que

$$v = (-\ln S) / 2T$$

Com essa fórmula, podemos estimar a taxa de evolução molecular de dois polipeptídios homólogos pelo (1) cálculo da proporção de sítios iguais nos dois, (2) determinação do logaritmo natural dessa proporção e depois (3) divisão pelo tempo evolutivo decorrido total. Em nosso exemplo,  $S = 0,75$  e  $2T = 300$  milhões de anos; assim  $v$  é  $[-\ln(0,75)] / 300 = 0,97$  substituição de aminoácidos por sítio a cada bilhão de anos.

Como discutido anteriormente, alguns dos sítios de aminoácidos diferentes em dois polipeptídios mudaram uma vez, outros mudaram duas vezes e ainda outros mudaram várias vezes durante o processo evolutivo. A quantidade  $2Tv$  é o número médio de substituições de aminoácidos ocorridas por sítio durante a evolução dos polipeptídios. Partindo do princípio de que as substituições de aminoácidos são aleatórias e independentes ao longo do tempo, podemos usar essa média para calcular a probabilidade de que um sítio tenha se modificado um número específico de vezes. O cálculo usa a fórmula para uma distribuição de probabilidade muito usada por cientistas. É denominada *distribuição de probabilidade de Poisson*. No contexto da evolução molecular, a fórmula de Poisson é

Probabilidade de que ocorram  $n$  modificações em um sítio de aminoácido =  $e^{-2Tv} (2Tv)^n / n!$

O número médio de substituições de aminoácidos ocorridas por sítio ( $2Tv$ ) aparece duas vezes nessa fórmula – como expoente do primeiro termo e como argumento da função de potência no segundo termo. Assim, é o parâmetro essencial da fórmula de Poisson.

Em nosso exemplo,  $2Tv$  é estimado por  $-\ln S = -\ln(0,75)$  como 0,29 substituição de aminoácido por sítio. Essa estimativa é um pouco maior que a proporção de aminoácidos diferentes nos dois polipeptídios ( $1 - S = 0,25$ ) porque leva em conta a possibilidade de que tenha havido múltiplas substituições em sítios de aminoácidos individuais. Dizemos que  $2Tv$  é o número de diferenças de aminoácidos corrigido pelo método de Poisson entre os dois polipeptídios.

Com uma estimativa de  $2Tv$ , podemos usar a fórmula de Poisson para calcular a probabilidade de que determinado sítio de aminoácido tenha se modificado exatamente uma vez, duas vezes, e assim por diante.

Probabilidade de uma modificação =  $e^{-2Tv} (2Tv) = 0,22$

Probabilidade de duas modificações =  $e^{-2Tv} (2Tv)^2 / 2 = 0,03$

A probabilidade de que não tenha havido modificações é

Probabilidade de nenhuma modificação =  $e^{-2Tv} = 0,75$

Nesse exemplo, a probabilidade de que tenha havido mais de duas modificações é desprezível. No entanto, se o parâmetro de Poisson  $2Tv$  fosse maior, poderia haver chance de múltiplas modificações. Por exemplo, se  $2Tv = 0,7$ , a probabilidade de três modificações em um sítio é 0,03, e a probabilidade de quatro modificações é 0,005.

Os procedimentos estatísticos análogos à correção de Poisson foram desenvolvidos para estimar taxas evolutivas a partir de comparações de sequências de DNA homólogas. Entretanto, esses procedimentos são mais complexos porque a identidade de um nucleotídeo em duas sequências de DNA não significa necessariamente que esse nucleotídeo permaneceu inalterado durante a evolução dessas sequências. Os métodos para lidar com essa questão são encontrados em textos especializados sobre o tema da evolução molecular.



# Respostas dos Problemas de Número Ímpar

## Capítulo 1

- 1.1 Mendel formulou a hipótese dos fatores transmissíveis – genes – para explicar a herança das características. Ele descobriu que os genes existem em diferentes formas, que agora chamamos de alelos. Cada organismo tem duas cópias de cada gene. Durante a reprodução, uma das cópias é incorporada aleatoriamente a cada gameta. Quando os gametas masculino e feminino se unem na fertilização, o número de cópias do gene volta a ser dois. Diferentes alelos podem coexistir em um organismo. Durante a produção de gametas eles se separam sem que tenham sido alterados pela coexistência.
- 1.3 As bases presentes no DNA são adenina, timina, guanina e citosina; as bases presentes no RNA são adenina, uracila, guanina e citosina. O açúcar do DNA é a desoxirribose; o açúcar do RNA é a ribose.
- 1.5 TAACGGCAG.
- 1.7 GAACGGUCT.
- 1.9 Às vezes o DNA é sintetizado a partir do RNA em um processo denominado transcrição reversa. Esse processo tem um papel importante nos ciclos vitais de alguns vírus.
- 1.11 As duas formas mutantes do gene da  $\beta$ -globina são denominadas alelos. Como nenhum dos alelos mutantes pode especificar um polipeptídeo “normal”, um indivíduo com as duas formas mutantes provavelmente teria anemia.

## Capítulo 2

- 2.1 Os açúcares combinam-se para formar carboidratos; os aminoácidos combinam-se para formar proteínas.
- 2.3 Em uma célula eucariótica os muitos cromossomos estão contidos em uma estrutura limitada por membrana denominada núcleo; os cromossomos das células procarióticas não estão contidos em

um compartimento subcelular especial. As células eucarióticas geralmente têm um sistema interno de membranas bem desenvolvido, e também têm organelas subcelulares delimitadas por membranas, como mitocôndrias e cloroplastos; as células procarióticas geralmente não têm sistema de membranas internas (embora algumas tenham) nem organelas delimitadas por membranas.

- 2.5 Os cromossomos procarióticos geralmente são menores que os cromossomos eucarióticos; além disso, os cromossomos procarióticos são circulares, enquanto os eucarióticos são lineares. Por exemplo, o cromossomo circular de *E. coli*, um procarioto, tem cerca de 1,4 mm de circunferência. Em contraposição, um cromossomo humano linear pode ter de 10 a 30 cm de comprimento. Os cromossomos procarióticos também têm uma composição comparativamente simples: DNA, alguma quantidade de RNA e alguma quantidade de proteína. Os cromossomos eucarióticos são mais complexos: DNA, alguma quantidade de RNA e grande quantidade de proteína.
- 2.7 A intérfase geralmente é mais demorada que a fase M. Durante a intérfase, é preciso que haja síntese de DNA para replicação de todos os cromossomos. Também é necessário que haja síntese de outros materiais no preparo para a divisão celular que está por vir.
- 2.9 (1) Anáfase: (f), (h); (2) metáfase: (e), (i); (3) prófase: (b), (c), (d); (4) telófase: (a), (g).
- 2.11 Não é esperado o pareamento dos cromossomos 11 e 16 durante a meiose; esses cromossomos são heterólogos, não homólogos.
- 2.13 O *crossing over* ocorre depois da duplicação dos cromossomos nas células em meiose.
- 2.15 A disjunção cromossômica ocorre durante a anáfase I. A disjunção das cromátides ocorre durante a anáfase II.
- 2.17 Em eucariotos, aparentemente não há nítida relação entre o tamanho do genoma e o número



de genes. Por exemplo, os seres humanos, com 3,2 bilhões de pares de bases de DNA genômico, têm cerca de 20.500 genes, e as plantas *Arabidopsis*, com cerca de 150 milhões de pares de bases de DNA genômico, têm quase o mesmo número de genes que os seres humanos. No entanto, em procariotos há correlação estreita entre o número de genes e o tamanho do genoma, provavelmente porque a quantidade de DNA não gênico é muito pequena.

- 2.19 É um tanto surpreendente o fato de que os cromossomos de levedura sejam, em média, menores que os cromossomos de *E. coli* porque, em regra, os cromossomos eucarióticos são maiores que os cromossomos procarióticos. A levedura é uma exceção porque seu genoma – que tem quase o triplo do tamanho do genoma de *E. coli* – é distribuído em 16 cromossomos.
- 2.21 Um dos núcleos do grão de pólen funde-se ao núcleo da oosfera no gametófito feminino e forma o zigoto, que dá origem a um embrião e, por fim, a um esporófito. O outro núcleo do grão de pólen geneticamente funcional funde-se a dois núcleos no gametófito feminino e forma um núcleo triploide, que dá origem a um tecido triploide, o endosperma; esse tecido nutre o embrião do vegetal em desenvolvimento.
- 2.23 (a) 5, (b) 5, (c) 15, (d) 10.

### Capítulo 3

- 3.1 (a) Todas altas; (b) 3/4 altas, 1/4 anã; (c) todas altas; (d) 1/2 alta, 1/2 anã.
- 3.3 Os dados sugerem que a cor da pelagem é controlada por um único gene com dois alelos, *C* (cinza) e *c* (albino), e que *C* é dominante em relação a *c*. Com base nessa hipótese, os cruzamentos são: cinza (*CC*) × albino (*cc*) →  $F_1$  cinza (*Cc*);  $F_1 \times F_1 \rightarrow 3/4$  cinza (1 *CC*: 2 *Cc*), 1/4 albino (*cc*). Os resultados esperados na  $F_2$  são 203 cinza e 67 albinos. Para comparar os resultados observados e esperados, calcule  $\chi^2$  com um grau de liberdade:  $(198 - 203)^2/203 + (72 - 67)^2/67 = 0,496$ , que não é relevante no nível de 5%. Portanto, os resultados estão de acordo com a hipótese.
- 3.5 (a) Quadriculada, vermelha (*CC BB*) × lisa, castanha (*cc bb*) →  $F_1$  toda quadriculada, vermelha (*Cc Bb*); (b) Prole  $F_2$ : 9/16 quadriculada, vermelha (*C- B-*), 3/16 lisa, vermelha (*cc B-*), 3/16 quadriculada, castanha (*C- bb*), 1/16 lisa, castanha (*cc bb*).
- 3.7 Na prole  $F_2$  com pelagem longa e preta, a proporção genotípica é 1 *BB RR*: 2 *Bb RR*: 2 *Bb Rr*: 4 *Bb rr*; assim, 1/9 dos coelhos com pelagem longa e preta é homozigoto para os dois genes.

### 3.9

| Gametas da $F_1$              | Genótipos da $F_2$         | Fenótipos da $F_2$                     |
|-------------------------------|----------------------------|----------------------------------------|
| (a) 2                         | 3                          | 2                                      |
| (b) $2 \times 2 = 4$          | $3 \times 3 = 9$           | $2 \times 2 = 4$                       |
| (c) $2 \times 2 \times 2 = 8$ | $3 \times 3 \times 3 = 27$ | $2 \times 2 \times 2 = 8$              |
| (d) $2^n$                     | $3^n$                      | $2^n$ , em que $n$ é o número de genes |

- 3.11 (a) 1, rejeição; (b) 2, rejeição; (c) 3, aceitação; (d) 3, aceitação.
- 3.13  $\chi^2 = (30 - 25)^2/25 + (20 - 25)^2/25 = 2$ , que é menor que 3,84, o valor de 5% crítico para a estatística de qui-quadrado com um grau de liberdade; consequentemente, a proporção de segregação observada está de acordo com a proporção esperada de 1:1.
- 3.15 Metade dos filhos de casamentos *Aa* × *aa* teria albinismo. Em uma família com três filhos, a chance de que um não seja afetado e dois sejam afetados é  $3 \times (1/2)^1 \times (1/2)^2 = 3/8$ .
- 3.17 Homem (*Cc ff*) × mulher (*cc Ff*). (a) *cc ff*,  $(1/2) \times (1/2) = 1/4$ ; (b) *Cc ff*,  $(1/2) \times (1/2) = 1/4$ ; (c) *cc Ff*,  $(1/2) \times (1/2) = 1/4$ ; (d) *Cc Ff*,  $(1/2) \times (1/2) = 1/4$ .
- 3.19  $(1/2)^3 = 1/8$ .
- 3.21  $(20/64) + (15/64) + (6/64) + (1/64) = 42/64$ .
- 3.23 (a)  $(1/2) \times (1/4) = 1/8$ ; (b)  $(1/2) \times (1/2) \times (1/4) = 1/16$ ; (c)  $(2/3) \times (1/4) = 1/6$ ; (d)  $(2/3) \times (1/2) \times (1/2) \times (1/4) = 1/24$ .
- 3.25 No casamento III-1 × III-2, a chance de ter um filho afetado é de 1/2. No casamento IV-2 × IV-3, a chance é igual a zero.
- 3.27 1/2.
- 3.29 Aparentemente, o pesquisador obteve uma proporção não mendeliana porque estudou somente famílias que têm pelo menos uma criança com albinismo. Nessas famílias, tanto o pai quanto a mãe são heterozigotos para o alelo mutante causador do albinismo. Entretanto, outros casais na população também podem ser heterozigotos para esse alelo, mas, por puro acaso, não tiveram filhos com albinismo. Se um homem e uma mulher são portadores heterozigotos do alelo mutante, a chance de que um filho não tenha albinismo é de 3/4. Portanto, a chance de que quatro filhos não tenham albinismo é  $(3/4)^4 = 0,316$ . Em toda a população de famílias nas quais pai e mãe heterozigotos tiveram um total de quatro filhos, o número médio de crianças afetadas é 1. Entre as famílias nas quais pai e mãe heterozigotos tiveram pelo menos uma criança afetada em um total de quatro filhos, a média tem de ser maior que 1. Para calcularmos essa *média condicional*, indiquemos como  $x$  o número de filhos com albinismo e como  $P(x)$ ,



a probabilidade de que exatamente  $x$  dos quatro filhos tenham albinismo. Portanto, o número médio de filhos afetados nas famílias em que pelo menos um dos quatro filhos é afetado – isto é, a média condicional – é  $\sum xP(x)/[1 - P(0)]$ , no qual a soma começa com  $x = 1$  e termina com  $x = 4$ . Iniciamos a soma com  $x = 1$  porque é preciso excluir os casos em que nenhum dos quatro filhos é afetado. O divisor  $[1 - P(0)]$  é a probabilidade de que o casal tenha pelo menos um filho afetado entre os quatro. Agora  $P(0) = 0,316$  e  $\sum xP(x) = 1$ . Portanto, a média que procuramos é simplesmente  $1/(1 - 0,316) = 1,46$ . Se, no subgrupo de famílias com pelo menos um filho afetado, o número médio de filhos afetados for 1,46, então o número médio de filhos não afetados será  $4 - 1,46 = 2,54$ . Assim, a proporção esperada de filhos não afetados e afetados nessas famílias é de 2,54:1,56 ou 1,74:1, que foi o observado pelo pesquisador.

## Capítulo 4

### 4.1 M e MN.

#### 4.3

| Pais                                            | Prole                                                       |
|-------------------------------------------------|-------------------------------------------------------------|
| (a) amarelos $\times$ amarelo                   | 2 amarelos: 1 ventre claro                                  |
| (b) amarelo $\times$ ventre claro               | 2 amarelos: 1 ventre claro:<br>1 preto e castanho-amarelado |
| (c) preto e castanho-amarelado $\times$ amarelo | 2 amarelos: 1 preto e castanho-amarelado: 1 preto           |
| (d) ventre claro $\times$ ventre claro          | todos com ventre claro                                      |
| (e) ventre claro $\times$ amarelo               | 1 amarelo: 1 ventre claro                                   |
| (f) agouti $\times$ preto e castanho-amarelado  | 1 agouti: 1 preto e castanho-amarelado                      |
| (g) preto e castanho-amarelado $\times$ preto   | 1 preto e castanho-amarelado:<br>1 preto                    |
| (h) amarelo $\times$ agouti                     | 1 amarelo: 1 ventre claro                                   |
| (i) amarelo $\times$ amarelo                    | 2 amarelos: 1 ventre claro                                  |

4.5 (a) todos AB; (b) 1 A: 1 B; (c) 1 A: 1 B: 1 AB: 1 O; (d) 1 A: 1 O.

4.7 Não. A mulher é  $I^A I^B$ . Um homem poderia ser  $I^A I^A$  ou  $I^A i$ ; o outro poderia ser  $I^B I^B$  ou  $I^B i$ . Ante a incerteza acerca do genótipo de cada homem, qualquer um deles poderia ser o pai da criança.

4.9 A mulher é  $ii L^M L^M$ ; o homem é  $I^A I^B L^M L^N$ ; os tipos sanguíneos das crianças serão A e M, A e MN, B e M, e B e MN, todos em igual probabilidade.

4.11 Os indivíduos III-4 e III-5 são necessariamente homozigotos para mutações recessivas em diferentes genes; ou seja, um é  $aa BB$  e o outro é  $AA bb$ ; nenhum dos filhos é surdo porque todos são heterozigotos para os dois genes ( $Aa Bb$ ).

4.13 Não. O teste de alelismo não pode ser feito com mutações dominantes.

4.15 A mãe é  $Bb$  e o pai é  $bb$ . A chance de que uma filha seja  $Bb$  é  $1/2$ . (a) A chance de que a filha tenha um filho calvo é  $(1/2) \times (1/2) = 1/4$ . (b) A chance de que a filha tenha uma filha calva é igual a zero.

4.17 (a)  $3/4$  com crista noz,  $1/4$  com crista rosa; (b)  $1/2$  com crista noz,  $1/2$  com crista ervilha; (c)  $3/8$  com crista noz,  $3/8$  com crista rosa,  $1/8$  com crista ervilha,  $1/8$  com crista simples; (d)  $1/2$  com crista rosa,  $1/2$  com crista simples.

4.19  $12/16$  com fruto branco,  $3/16$  com fruto amarelo,  $1/16$  com fruto verde.

4.21  $9/16$  com olhos vermelho-escuros (tipo selvagem),  $3/16$  com olhos roxo-acastanhados,  $3/16$  com olhos vermelho-brilhantes,  $1/16$  com olhos brancos.

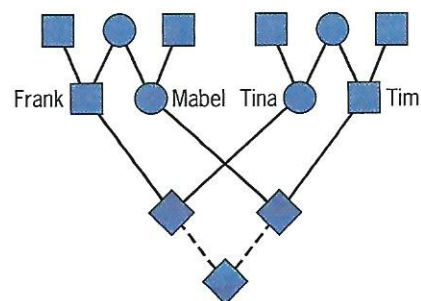
4.23 9 com pelagem preta: 39 com pelagem cinza: 16 com pelagem branca.

4.25 (a) roxa  $\times$  vermelha; (b) proporção com sementes brancas ( $aa$ ) =  $1/4$ ; (c) proporção de prole vermelha ( $A- B- C- dd$ ) =  $(3/4)(3/4)(3/4)(1/2) = 27/128$ , proporção de prole branca ( $aa$ ) =  $1/4$  =  $32/128$ , proporção de prole azul ( $A- B- cc Dd$ ) =  $(3/4)(3/4)(1/4)(1/2) = 9/128$ .

4.27 (a) Como a segregação da  $F_2$  é de aproximadamente 9 vermelhas: 7 brancas, a cor das flores se deve à epistasia entre dois genes de distribuição independente: vermelho =  $A- B-$  e branco =  $aa B-$ ,  $A- bb$  ou  $aa bb$ . (b) Precursor incolor  $-A \rightarrow$  produto incolor  $- B \rightarrow$  pigmento vermelho.

4.29  $F_A = (1/2)^5 = 1/32$ ;  $F_B = 2 \times (1/2)^6 = 1/32$ ;  $F_C = 2 \times (1/2)^7 = 1/64$ .

4.31 O heredograma é o seguinte:



O coeficiente de relação entre a prole dos dois casais é obtido pelo cálculo do coeficiente de endogamia do filho imaginário de um casamento desses filhos e multiplicação desse número por 2:  $[(1/2)^5 \times 2] \times 2 = 1/8$ .

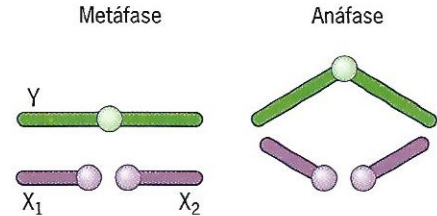
4.33 O comprimento médio da espiga do milho obtido por cruzamento aleatório é de 24 cm, e o do milho de uma geração de autofertilização é de 20 cm. O



coeficiente de endogamia da prole de uma geração de autofertilização é  $1/2$ , e o coeficiente de endogamia da prole de duas gerações de autofertilização é  $(1/2)(1 + 1/2) = 3/4$ . Há previsão de declínio linear do comprimento médio da espiga ( $Y$ ) de acordo com a equação  $Y = 24 - bF_1$ , em que  $b$  é o coeficiente angular da reta. Pode-se calcular o valor de  $b$  a partir dos dois valores de  $Y$  dados. A diferença entre esses dois valores (4 cm) corresponde a um aumento de  $F$  de 0 para  $1/2$ . Assim,  $b = 4(1/2) = 8$  cm, e para  $F = 3/4$ , o comprimento médio da espiga previsto é  $Y = 24 - 8 \times (3/4) = 18$  cm.

## Capítulo 5

- 5.1 O espermatozoide determinante do sexo masculino tem um cromossomo Y; o espermatozoide determinante do sexo feminino tem um cromossomo X.
- 5.3 Todas as fêmeas terão corpo verde, e todos os machos terão corpo rosado.
- 5.5 XX é fêmea, XY é macho, XXY é fêmea, XXX é fêmea (mas dificilmente viável), XO é macho (mas estéril).
- 5.7 Não. A discromatopsia é causada por uma mutação ligada ao X. O cromossomo X do filho é oriundo da mãe, não do pai.
- 5.9 O risco para a criança é  $P$  (mãe  $C/c$ )  $\times$   $P$  (mãe transmitir  $c$ )  $\times$   $P$  (criança de sexo masculino) =  $(1/2) \times (1/2) \times (1/2) = 1/8$ ; se o casal já teve um filho com discromatopsia,  $P$  (mãe  $C/c$ ) = 1, e o risco de cada filho subsequente é  $1/4$ .
- 5.11 Cada uma das filhas com olhos vermelhos foi produzida necessariamente pela união de um ovócito  $X(v)$   $X(v)$  com um espermatozoide que tem  $Y$ . Os ovócitos duplo- $X$  originaram-se obrigatoriamente da não disjunção do cromossomo  $X$  durante a ovocitogênese na mãe. Entretanto, não podemos determinar se a não disjunção ocorreu na primeira ou na segunda divisão meiótica.
- 5.13 Cada uma das raras filhas com olhos brancos foi produzida necessariamente pela união de um ovócito  $X(w)$   $X(w)$  com um espermatozoide que tem  $Y$ . Os raros ovócitos duplo- $X$  originaram-se obrigatoriamente da não disjunção do cromossomo  $X$  durante a segunda divisão meiótica na mãe.
- 5.15 Feminino.
- 5.17 Homem.
- 5.19 (a) Feminino; (b) intersexo; (c) intersexo; (d) masculino; (e) feminino; (f) masculino.
- 5.21 A *Drosophila* não obtém compensação de dose por inativação de um dos cromossomos  $X$  em fêmeas.
- 5.23 Como o centrômero está na extremidade de cada pequeno cromossomo  $X$ , mas no meio do cromossomo  $Y$  maior,  $X_1$  e  $X_2$  emparelham-se no centrômero de  $Y$  durante a metáfase. Então, durante a anáfase, os dois cromossomos  $X$  sofrem disjunção e separam-se do cromossomo  $Y$ .



- 5.25 A cor dos olhos de canários é determinada por um gene no cromossomo  $Z$ , que está presente em duas cópias nos machos e uma cópia nas fêmeas. O alelo para cor rosa ao nascer ( $p$ ) é recessivo em relação ao alelo para cor preta ao nascer ( $P$ ). Não há gene para cor dos olhos no outro cromossomo sexual ( $W$ ), presente em uma cópia nas fêmeas e ausente nos machos. As aves parentais têm genótipo  $p/W$  (fêmeas canela) e  $P/P$  (machos verdes). Os machos da  $F_1$  tinham genótipo  $p/P$  (com olhos pretos ao nascer). Quando os machos da  $F_1$  foram cruzados com fêmeas verdes (genótipo  $P/W$ ), produziram prole  $F_2$  dividida em três categorias: machos com olhos pretos ao nascer ( $P/-$ , metade da prole), fêmeas com olhos pretos ao nascer ( $P/W$ , um quarto da prole), e fêmeas com olhos rosa ao nascer ( $p/W$ , um quarto da prole). Quando os machos da  $F_1$  foram cruzados com fêmeas canela (genótipo  $p/W$ ), produziram prole  $F_2$  dividida em quatro categorias de frequências iguais: machos com olhos pretos ao nascer (genótipo  $P/p$ ), machos com olhos rosa ao nascer (genótipo  $p/p$ ), fêmeas com olhos pretos ao nascer (genótipo  $P/W$ ) e fêmeas com olhos rosa ao nascer (genótipo  $p/W$ ).

## Capítulo 6

- 6.1 Uso de uma das técnicas de bandejamento.
- 6.3 Em alotetraploides, pode haver pareamento de cada membro dos diferentes conjuntos de cromossomos com um homólogo durante a prófase I, seguido por disjunção durante a anáfase I. Em triploides, a disjunção é irregular porque os cromossomos homólogos associam-se durante a prófase I formando bivalentes e univalentes ou formando trivalentes.
- 6.5 O vegetal fértil é um alotetraploide com 7 pares de cromossomos da espécie A e 9 pares de cromossomos da espécie B; o número total de cromossomos é  $(2 \times 7) + (2 \times 9) = 32$ .
- 6.7 XX é feminino, XY é masculino, XO é feminino (mas estéril), XXX é feminino, XXY é masculino (mas estéril), XYY é masculino.

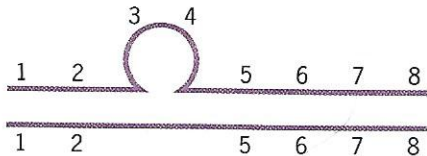


6.9 A mosca é um ginandromorfo, ou seja, um mosaico sexual. O tecido amarelo é  $X(y)/O$  e o tecido cinza é  $X(y)/X(+)$ . Esse mosaicismo foi causado pela perda do cromossomo X que tinha o alelo selvagem, provavelmente durante uma das divisões iniciais da clivagem embrionária.

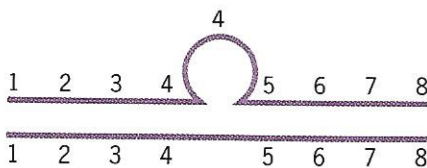
6.11 A não disjunção ocorreu obrigatoriamente na mãe. A mulher com discromatopsia e síndrome de Turner foi produzida pela união de um espermatozoide com X, que tinha o alelo mutante para discromatopsia, e um ovócito nulo-X.

6.13 Homens XYY teriam mais filhos com anormalidades dos cromossomos sexuais em razão da disjunção irregular de seus três cromossomos sexuais durante a meiose. Essa disjunção irregular produz uma variedade de gametas aneuploides, que inclui as constituições cromossômicas XY, YY, XYY e ausência de cromossomo sexual.

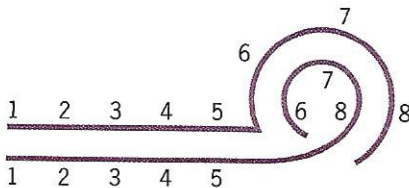
6.15 (a) Deleção:



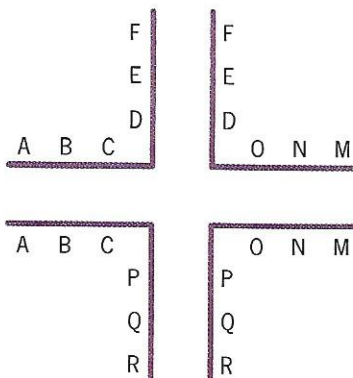
(b) Duplicação:



(c) Inversão terminal:



6.17

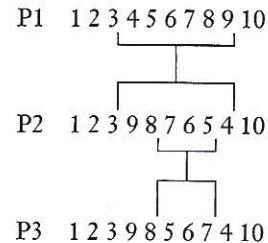


6.19 O menino tem uma translocação entre o cromossomo 21 e outro cromossomo, por exemplo, o cromossomo 14. Ele também tem um cromossomo 21

normal e um cromossomo 14 normal. A irmã do menino tem a translocação, um cromossomo 14 normal e duas cópias normais do cromossomo 21.

6.21 Todas as filhas terão corpo amarelo, e todos os filhos terão olhos brancos.

6.23 As três populações estão relacionadas por uma série de inversões.



6.25 A mãe é heterozigota para uma translocação recíproca entre os braços longos dos cromossomos grande e pequeno; um trecho do braço longo do cromossomo grande foi quebrado e fixado ao braço longo do cromossomo pequeno. A criança herdou o cromossomo grande com rearranjo e o cromossomo pequeno normal da mãe. Portanto, como o cromossomo grande com rearranjo tem deficiência de alguns genes, a criança é hipoploide.

6.27 Os filhos terão olhos vermelho-brilhantes porque herdarão do pai o cromossomo Y com o alelo  $bw^+$ . As filhas terão olhos brancos porque herdarão um cromossomo X do pai.

6.29 Os zigotos XX darão origem a homens porque um de seus cromossomos X tem o gene *SRY* translocado do cromossomo Y. Os zigotos XY darão origem a mulheres porque o cromossomo Y perdeu o gene *SRY*.

## Capítulo 7

7.1 Se Mendel soubesse da existência dos cromossomos, teria percebido que o número de fatores de que determinam as características é maior que o número de cromossomos e concluído que alguns fatores têm de estar ligados no mesmo cromossomo. Portanto, ele teria reformulado o princípio da distribuição independente e afirmaria que a herança de fatores em diferentes cromossomos (ou distantes no mesmo cromossomo) é independente.

7.3 Não. Os genes *a* e *d* podem estar muito distantes no mesmo cromossomo – tão distantes que se recombina livremente, isto é, 50% das vezes.

7.5 Sim, se eles estiverem muito distantes.

7.7 (a) Cruzamento:  $a^+ b^+ / a^+ b^+ \times a b / a b$ . Gametas:  $a^+ b^+$  de um genitor,  $a b$  do outro.  $F_1$ :  $a^+ b^+ / a b$ . (b) 40%,  $a^+ b^+$ ; 40%,  $a b$ ; 10%,  $a^+ b$ ; 10%,  $a b^+$ . (c)  $F_2$  do cruzamento-teste: 40%,  $a^+ b^+ / a b$ ; 40%,  $a b / a b$ ;



10%,  $a^+ b/a b$ ; 10%,  $a b^+/a b$ . (d) Fase de ligação de acoplamento. (e)  $F_2$  do inter cruzamento:

|               | Espermatozoides          |                      |                       |                       |
|---------------|--------------------------|----------------------|-----------------------|-----------------------|
|               | 40% $a^+ b^+$            | 40% $a b$            | 10% $a^+ b$           | 10% $a b^+$           |
| 40% $a^+ b^+$ | 16%<br>$a^+ b^+/a^+ b^+$ | 16%<br>$a^+ b^+/a b$ | 4%<br>$a^+ b^+/a^+ b$ | 4%<br>$a^+ b^+/a b^+$ |
| 40% $a b$     | 16%<br>$a b/a^+ b^+$     | 16%<br>$a b/a b$     | 4%<br>$a b/a^+ b$     | 4%<br>$a b/a b^+$     |
| 10% $a^+ b$   | 4%<br>$a^+ b/a^+ b^+$    | 4%<br>$a^+ b/a b$    | 1%<br>$a^+ b/a^+ b$   | 1%<br>$a^+ b/a b^+$   |
| 10% $a b^+$   | 4%<br>$a b^+/a^+ b^+$    | 4%<br>$a b^+/a b$    | 1%<br>$a b^+/a^+ b$   | 1%<br>$a b^+/a b^+$   |

Resumo dos fenótipos:

|             |     |           |     |
|-------------|-----|-----------|-----|
| $a^+ e b^+$ | 66% | $a e b^+$ | 9%  |
| $a^+ e b$   | 9%  | $a e b$   | 16% |

7.9 Heterozigotos em acoplamento  $a^+ b^+/a b$  produziram os seguintes gametas: 30%,  $a^+ b^+$ ; 30%,  $a b$ ; 20%,  $a^+ b$ ; 20%,  $a b^+$ ; heterozigotos em repulsão  $a^+ b/a b^+$  produziram os seguintes gametas: 30%,  $a^+ b$ ; 30%,  $a b^+$ ; 20%,  $a^+ b^+$ ; 20%,  $a b$ . Em cada caso, as frequências da prole do cruzamento-teste corresponderiam às frequências dos gametas.

7.11 Sim. Frequência de recombinação =  $(24 + 26)/(126 + 24 + 26 + 124) = 0,167$ . Cruzamento:

| $b vg$ fêmea |   | $b vg$ macho |
|--------------|---|--------------|
| $b^+ vg^+$   | × | $b vg$       |
| $b vg$       | ↓ | $b^+ vg^+$   |
| $b^+ vg^+$   | ↓ | $b vg$       |
| 126          |   | 24           |
| 124          |   | 26           |

7.13 Sim. A frequência de recombinação é estimada pela frequência de prole preta entre a prole colorida:  $34/(66 + 34) = 0,34$ . Cruzamento:

| $C b$    |   | $c B$ |
|----------|---|-------|
| $C b$    | × | $c B$ |
| $C b$    | ↓ | $c b$ |
| $c B$    | ↓ | $C b$ |
| 66       |   | 34    |
| castanha |   | preta |
| 100      |   |       |

7.15 (a) As fêmeas da  $F_1$ , que são  $sr e^+/sr^+ e$ , produzem quatro tipos de gametas: 46%,  $sr e^+$ ; 46%,  $sr^+ e$ ; 4%,  $sr e$ ; 4%,  $sr^+ e^+$ . (b) Os machos da  $F_1$ , que têm o mesmo genótipo que as fêmeas da  $F_1$ , produzem dois tipos de gametas: 50%,  $sr e^+$ ; 50%,  $sr^+ e$ ; lembre-se de que não há *crossing over* em machos de *Drosophila*. (c) 46%, listrados e com corpo cinza; 46%, sem listras e com corpo ébano; 4%, listrados e com corpo ébano; 4%, sem listras e com corpo cinza. (d)

A prole do inter cruzamento é obtida a partir da seguinte tabela.

|            | Espermatozoides   |                   |
|------------|-------------------|-------------------|
|            | $sr e^+$          | $sr^+ e$          |
|            | 0,50              | 0,50              |
| $sr e^+$   | $sr e^+/sr e^+$   | $sr e^+/sr^+ e$   |
| 0,46       | 0,23              | 0,23              |
| $sr^+ e$   | $sr^+ e/sr e^+$   | $sr^+ e/sr^+ e$   |
| 0,46       | 0,23              | 0,23              |
| $sr e$     | $sr e/sr e^+$     | $sr e/sr^+ e$     |
| 0,04       | 0,002             | 0,002             |
| $sr^+ e^+$ | $sr^+ e^+/sr e^+$ | $sr^+ e^+/sr^+ e$ |
| 0,04       | 0,002             | 0,002             |

7.17 (a) As fêmeas da  $F_1$ , que são  $cn vg^+/cn^+ vg$ , produzem quatro tipos de gametas: 45%,  $cn vg^+$ ; 45%,  $cn^+ vg$ ; 5%,  $cn^+ vg^+$ ; 5%,  $cn vg$ . (b) 45%, olhos cinabre, asas normais; 45%, olhos castanho-avermelhados, asas vestigiais; 5%, olhos castanho-avermelhados, asas normais; 5%, olhos cinabre, asas vestigiais.

7.19 Na enumeração adiante, as classes 1 e 2 são tipos parentais, as classes 3 e 4 são resultado de um único *crossing over* entre *Pl* e *Sm*, as classes 5 e 6 são resultado de um único *crossing over* entre *Sm* e *Py*, e as classes 7 e 8 são resultado de um *crossing over* duplo, com uma das permutas entre *Pl* e *Sm* e a outra entre *Sm* e *Py*.

| Classe | Fenótipos              | (a) Frequência sem inter-ferência | (b) Frequência com inter-ferência completa |
|--------|------------------------|-----------------------------------|--------------------------------------------|
| 1      | roxo, salmão, pigmeu   | 0,405                             | 0,40                                       |
| 2      | verde, amarelo, normal | 0,405                             | 0,40                                       |
| 3      | roxo, amarelo, normal  | 0,045                             | 0,05                                       |
| 4      | verde, salmão, pigmeu  | 0,045                             | 0,05                                       |
| 5      | roxo, salmão, normal   | 0,045                             | 0,05                                       |
| 6      | verde, amarelo, pigmeu | 0,045                             | 0,05                                       |
| 7      | roxo, amarelo, pigmeu  | 0,005                             | 0                                          |
| 8      | verde, salmão, normal  | 0,005                             | 0                                          |

7.21 As classes de *crossing over* duplo, que são as duas não observadas, determinam que a ordem do gene é  $y - w - ec$ . Assim, as fêmeas da  $F_1$  tinham o genótipo  $y w ec/++$ . A distância entre *y* e *w* é estimada pela frequência de recombinação entre esses dois genes:  $(8 + 7)/1.000 = 0,015$ ; da mesma maneira, a distância entre *w* e *ec* é  $(18 + 23)/1.000 = 0,041$ . Assim, o mapa genético desse segmento do cromossomo X é  $y-1,5 \text{ cM}-w-4,1 \text{ cM}-ec$ .

7.23 (a) Duas classes (os tipos parentais) são muito mais numerosas que as outras seis classes (tipos recombinantes); (b)  $st +/+ + ss e$ ; (c)  $st-ss-e$ ; (d)  $[(145 + 122) \times 1 + (18) \times 2]/1.000 = 30,3 \text{ cM}$ ; (e)  $(122 + 18)/1.000 = 14,0 \text{ cM}$ ; (f)  $(0,018)/(0,163 \times 0,140) = 0,789$ . (g) fêmeas  $st +/+ + ss e \times$  machos  $st ss e/st ss e \rightarrow 2$  classes parentais e 6 classes recombinantes.



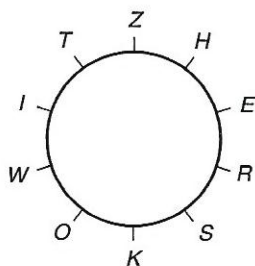
- 7.25 As fêmeas da  $F_1$  têm genótipo  $pn +/+ g$ . Entre os filhos machos, 40% serão recombinantes para os dois genes ligados ao X, e metade dos recombinantes terá os alelos selvagens desses genes. Assim, a frequência de filhos machos com olhos vermelho-escuros será de  $1/2 \times 40\% = 20\%$ .
- 7.27  $(P/2)^2$ .
- 7.29 Das classes parentais,  $++c$  e  $ab+$ , as fêmeas heterozigotas tinham, obrigatoriamente, o genótipo  $++c/ab+$ . As classes ausentes,  $+b+$  e  $a+c$ , que representariam *crossing overs* duplos, determinam que a ordem dos genes é  $b-a-c$ . A distância entre  $b$  e  $a$  é  $(96 + 110)/1.000 = 20,6$  cM, e entre  $a$  e  $c$  é  $(65 + 75)/1.000 = 14,0$  cM. Assim, o mapa genético é  $b-20,6$  cM- $a-14,0$  cM- $c$ .
- 7.31 II-1 tem o genótipo  $Ch/cH$ ; ou seja, ela é heterozigota em repulsão para os alelos para discromatopsia ( $c$ ) e hemofilia ( $h$ ). Nenhum dos seus filhos é recombinante para esses alelos.
- 7.33 A mulher é um heterozigoto em repulsão para os alelos para discromatopsia e hemofilia – ou seja, ela é  $Ch/cH$ . Se tiver um menino, a chance de que ele tenha hemofilia é de 0,5 e a chance de que tenha discromatopsia é de 0,5. Se especificarmos que o menino tem apenas um desses distúrbios, a chance de que tenha discromatopsia é de 0,45. A explicação é que a probabilidade de que o menino herde um cromossomo X não recombinante é de 0,9, e metade dos cromossomos X não recombinantes terá o alelo mutante para discromatopsia e a outra metade terá o alelo mutante para hemofilia. A chance de que o menino tenha os dois distúrbios é de 0,05, e a chance de que não tenha nenhum dos dois é de 0,05. A explicação é que a probabilidade de que o menino herde um cromossomo X recombinante é de 0,1, e metade dos cromossomos X recombinantes terá os dois alelos mutantes e a outra metade não terá nenhum alelo mutante.
- 7.35 Um *crossing over* duplo bifilamentar dentro da inversão; os pontos de permuta do *crossing over* duplo têm de estar entre os marcadores genéticos e os pontos de quebra da inversão.

## Capítulo 8

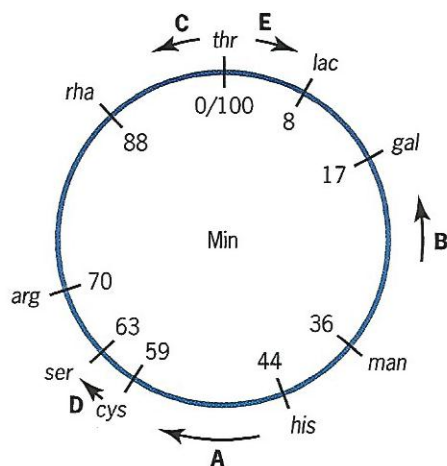
- 8.1 Os vírus reproduzem-se e transmitem os seus genes para a prole. Usam a energia das células hospedeiras e respondem a sinais ambientais e celulares da mesma forma que outros organismos vivos. Os vírus, porém, são parasitos obrigatórios; eles só se reproduzem em células hospedeiras apropriadas.
- 8.3 O bacteriófago T4 é um fago virulento. Ao infectar uma célula hospedeira, ele se reproduz e destrói a célula durante esse processo. O bacteriófago lambda pode se reproduzir e destruir a bactéria hospedeira – resposta lítica – da mesma maneira que o fago T4, ou pode inserir seu cromossomo no cromossomo da célula hospedeira e permanecer aí em estado latente – resposta lisogênica.
- 8.5 A inserção do cromossomo do fago  $\lambda$  no cromossomo da célula hospedeira é um processo de recombinação sítio-específico catalisado por uma enzima que reconhece sequências específicas nos cromossomos de  $\lambda$  e da *E. coli*. O *crossing over* entre cromossomos homólogos não é específico para sequências. Pode ocorrer em muitos sítios ao longo dos dois cromossomos.
- 8.7 As mutações  $a$ ,  $b$  e  $c$  estão próximas e na ordem  $b-a-c$  no cromossomo.
- 8.9 Faça dois experimentos: (1) verifique se o processo é sensível à DNase, e (2) verifique se há necessidade de contato celular para que ocorra o processo. A necessidade de contato celular pode ser testada por um experimento com tubo em U (Figura 8.9). Se o processo for sensível à DNase, é semelhante à transformação. Caso haja necessidade de contato celular, é semelhante à conjugação. Se não for sensível à DNase nem exigir contato celular, é semelhante à transdução.
- 8.11 (a) Os fatores  $F'$  são úteis nas análises genéticas em que são necessárias duas cópias de um gene na mesma célula, por exemplo, para determinar relações de dominância. (b) Os fatores  $F'$  são formados por excisão anormal de fatores  $F$  dos cromossomos de células Hfr (Figura 8.21). (c) Transferência conjugativa de um fator  $F'$  de uma célula doadora para uma célula receptora ( $F^-$ ).
- 8.13 Os elementos IS (sequências de inserção) são sequências de DNA curtas (800 a 1.400 pares de nucleotídeos) transponíveis, isto é, capazes de se deslocarem de uma posição para outra em um cromossomo ou de um cromossomo para outro. Os elementos IS medeiam a recombinação entre moléculas de DNA não homólogas – por exemplo, entre fatores  $F$  e cromossomos bacterianos.
- 8.15 Cotransdução é a transdução simultânea de dois marcadores genéticos diferentes para uma única célula receptora. Já que as partículas de bacteriófagos só comportam 1/100 a 1/50 do cromossomo bacteriano total, só podem ser cotransduzidos os marcadores que estão relativamente próximos. A frequência de cotransdução de dois marcadores quaisquer é inversamente proporcional à distância entre eles no cromossomo. Assim, essa frequência pode ser usada para estimar a distância de ligação gênica. É preciso preparar relações de ligação-cotransdução específicas para cada sistema fago-hospedeiro estudado.



8.17

8.19 *pro-pur-his.*

8.21



## Capítulo 9

9.1 (a) Os experimentos *in vivo* de Griffith mostraram a transformação em pneumococos. Não indicaram a base molecular do fenômeno de transformação. Avery e colaboradores fizeram experimentos *in vitro*, empregando análises bioquímicas para mostrar que a transformação era mediada por DNA. (b) Griffith mostrou que existia uma substância transformadora; Avery *et al.* definiram que era o DNA. (c) Os experimentos de Griffith não incluíram nenhuma tentativa de caracterizar a substância responsável pela transformação. Avery *et al.* isolaram o DNA na forma "pura" e mostraram que poderia mediar a transformação.

9.3 Demonstrou-se que o DNA purificado do tipo III era suficiente para transformar células tipo II. Isso ocorreu na ausência de células tipo III mortas.

9.5 (a) O objetivo era determinar se o material genético era DNA ou proteína. (b) Marcando o fósforo, um constituinte do DNA, e o enxofre, um constituinte da proteína, em um vírus, foi possível demonstrar que apenas o fósforo marcado era introduzido na célula hospedeira durante o ciclo reprodutivo do vírus. O DNA era suficiente para produzir novos fagos. (c) Portanto, o DNA, e não a proteína, é o material genético.

9.7 (a) O padrão em escada era conhecido a partir dos estudos de difração por raios X. Análises químicas mostraram a existência de uma relação 1:1 entre as bases orgânicas adenina e timina e entre

citosina e guanina. Eles também dispunham de dados físicos acerca do comprimento de cada espiral e do empilhamento de bases. (b) Watson e Crick desenvolveram o modelo de uma dupla hélice, com os filamentos rígidos de açúcar e fósforo formando espirais ao redor de um eixo, e ligações de hidrogênio unindo as bases complementares em pares de nucleotídeos.

9.9 (a) 400.000; (b) 20.000; (c) 400.000; (d) 68.000 nm.

9.11 Não. O RNA do TMV é unifilar. Assim, não é válida a relação estequiométrica existente nos pares de bases do DNA.

9.13 3'-CAGTACTG-5'.

9.15 (a) DNA bifilar; (b) DNA unifilar; (c) RNA unifilar.

9.17 O valor de  $T_m$  aumenta com o conteúdo de GC porque os pares de bases GC, unidos por três ligações de hidrogênio, são mais fortes que os pares de bases AT, unidos por duas ligações de hidrogênio.

9.19 (1) O nível do nucleossomo; o centro contendo um octâmero de histonas mais 146 pares de nucleotídeos de DNA dispostos em  $1\frac{3}{4}$  voltas de uma super-hélice (ver Figura 9.18), produzindo um corpo esférico com diâmetro aproximado de 11 nm; ou justapostos, uma fibra com diâmetro aproximado de 11 nm. (2) A fibra de 30 nm observada nos cromossomos mitóticos e meióticos condensados; ela parece ser formada por helicoidização ou dobramento da fibra do nucleossomo de 11 nm. (3) Os cromossomos mitóticos e meióticos altamente condensados (p. ex., cromossomos metafásicos); o dobramento ou a helicoidização acentuados mantidos por um "esqueleto" composto de proteínas cromossômicas não histônicas (ver Figura 9.22).

9.21 (a) 89,5°C. (b) Aproximadamente 39%.

9.23 A renaturação dos fragmentos de DNA satélite seria muito mais rápida que a dos fragmentos de DNA da banda principal. Nos DNA satélites de *D. virilis*, todos os três têm sequências de pares repetidos de heptanucleotídeos. Assim, praticamente todo fragmento unifilar com 40 nucleotídeos (médio) de um filamento terá uma sequência complementar (em parte) a cada fragmento unifilar do filamento complementar. Muitas das sequências de pares de nucleotídeos no DNA da banda principal serão sequências únicas (presentes apenas uma vez no genoma).

9.25 Intérfase. Os cromossomos são, em sua maior parte, metabolicamente inativos (apresentam baixa transcrição) durante os vários estágios da condensação na mitose e na meiose.



- 9.27 (a) As histonas foram altamente conservadas durante toda a evolução dos eucariotos. Uma importante função das histonas é empacotar o DNA em nucleossomos e fibras de cromatina. Como o DNA é constituído dos mesmos quatro nucleotídeos e tem a mesma estrutura básica em todos os eucariotos, seria de se esperar que as proteínas que desempenham um papel estrutural no empacotamento desse DNA fossem conservadas da mesma maneira. (b) As proteínas cromossômicas não histônicas apresentam a maior heterogeneidade na cromatina de diferentes tecidos e tipos celulares de um organismo. A composição histônica é basicamente igual em todos os tipos celulares de uma espécie – compatível com o papel das histonas no empacotamento de DNA em nucleossomos. As proteínas cromossômicas não histônicas abrangem proteínas que regulam a expressão gênica. Como os diferentes conjuntos de genes são transcritos em diferentes tipos celulares, seria de se esperar a heterogeneidade em algumas das proteínas cromossômicas não histônicas de diferentes tecidos.

## Capítulo 10

- 10.1 (a) Atividades de exonuclease, tanto  $3' \rightarrow 5'$  quanto  $5' \rightarrow 3'$ . (b) A exonuclease  $3' \rightarrow 5'$  “revisa” o filamento nascente de DNA durante a síntese. Se houver um erro de pareamento de bases na extremidade  $3'$ -OH do iniciador, a exonuclease  $3' \rightarrow 5'$  retira o nucleotídeo terminal incorreto antes que seja retomada a polimerização. A exonuclease  $5' \rightarrow 3'$  é responsável pela remoção de iniciadores de RNA durante a replicação do DNA e atua em vias participantes do reparo do DNA lesado (Capítulo 13). (c) Sim, as duas atividades de exonuclease parecem ser muito importantes. Sem a atividade de revisão  $3' \rightarrow 5'$  durante a replicação, a frequência de mutações seria intolerável. A atividade de exonuclease  $5' \rightarrow 3'$  é essencial para a sobrevivência da célula. Mutações condicionais que modificam a atividade de exonuclease  $5' \rightarrow 3'$  da DNA polimerase I são letais para a célula se houver inatividade da exonuclease.
- 10.3 O  $^{15}\text{N}$  contém oito nêutrons em vez dos sete nêutrons existentes no isótopo normal do nitrogênio,  $^{14}\text{N}$ . Portanto, a massa atômica aproximada de  $^{15}\text{N}$  é 15, enquanto a massa aproximada de  $^{14}\text{N}$  é 14. Essa diferença significa que as purinas e pirimidinas com  $^{15}\text{N}$  têm maior densidade (massa por unidade de volume) que aquelas com  $^{14}\text{N}$ . A centrifugação de equilíbrio por gradiente de densidade em  $\text{CsCl } 6M$  separa o DNA ou outras macromoléculas de acordo com a densidade, e o DNA de *E. coli*, por exemplo, que contém  $^{15}\text{N}$  tem densidade de  $1,724 \text{ g/cm}^3$ , enquanto o DNA de *E. coli* que contém  $^{14}\text{N}$  tem densidade de  $1,710 \text{ g/cm}^3$ .
- 10.5 Se o DNA nascente fosse marcado por exposição à  $^3\text{H}$ -timidina por períodos muito curtos, na replicação contínua o marcador seria incorporado a moléculas de DNA do tamanho do cromossomo, enquanto na replicação descontínua o marcador apareceria primeiro em pequenos trechos do DNA nascente (antes da ligação covalente, catalisada por DNA ligase).
- 10.7
- 
- 10.9 Esse tipo de autorradiografia indicaria replicação unidirecional do DNA. À medida que há diluição gradual da  $^3\text{H}$ -timidina intracelular depois da transferência para meio não radioativo, é cada vez menor a quantidade de  $^3\text{H}$ -timidina incorporada ao DNA em cada forquilha de replicação. Isso produz autorradiografias com densidade decrescente dos grãos nas caudas em cada ponto de crescimento. Como essas caudas só aparecem em uma extremidade de cada linha, a replicação é obrigatoriamente unidirecional. A replicação bidirecional produziria essas caudas nas duas extremidades da linha na autorradiografia (Figura 10.31).
- 10.11 As evidências atuais sugerem que as polimerases  $\alpha$ ,  $\delta$  e/ou  $\epsilon$  são necessárias para a replicação do DNA nuclear. Acredita-se que as polimerases  $\delta$  e/ou  $\epsilon$  catalisem a síntese contínua do filamento líder, e que a polimerase  $\alpha$  atue como primase na síntese descontínua do filamento atrasado. A polimerase  $\gamma$  catalisa a replicação dos cromossomos organelares. As polimerases  $\beta$ ,  $\zeta$ ,  $\eta$ ,  $\theta$ ,  $\iota$ ,  $\kappa$ ,  $\lambda$ ,  $\mu$ ,  $\sigma$ ,  $\phi$  e Rev1 atuam em várias vias de reparo do DNA (Capítulo 13).
- 10.13 Não haverá DNA na posição “leve”; 12,5% (2 das 16 moléculas de DNA) terão densidade “híbrida”; e 87,5% (14 das 16 moléculas de DNA) estarão na posição “pesada”.
- 10.15 (a) DNA girase; (b) primase; (c) atividade de exonuclease  $5' \rightarrow 3'$  da DNA polimerase I; (d) atividade de polimerase  $5' \rightarrow 3'$  da DNA polimerase III; (e) atividade de exonuclease  $3' \rightarrow 5'$  da DNA polimerase III.
- 10.17 Em eucariotos, a taxa de síntese de DNA em cada forquilha de replicação é de aproximadamente 2.500 a 3.000 pares de nucleotídeos por minuto. Com frequência, grandes cromossomos eucarióticos contêm de  $10^7$  a  $10^8$  pares de nucleotídeos.



Uma única forquilha de replicação não seria capaz de replicar o DNA gigante em um desses cromossomos grandes com velocidade suficiente para possibilitar os tempos de geração celular observados.

- 10.19** Não. Linhagens de *E. coli* com mutações *polA* que eliminam a atividade de exonuclease 3' → 5' da DNA polimerase I terão alta taxa de mutação.
- 10.21** (a) A replicação por círculo rolante começa quando uma endonuclease cliva um filamento de uma dupla hélice de DNA circular. Essa clivagem produz um grupo 3'-OH livre em uma extremidade do filamento cortado, possibilitando que este atue como um iniciador. (b) A síntese descontínua do filamento atrasado requer a nova iniciação de cada fragmento de Okazaki, o que exige atividade da DNA primase.
- 10.23** A DNA helicase desenrola a dupla hélice de DNA, e a proteína de ligação ao DNA unifilamentar recobre os filamentos desenrolados, mantendo-os estendidos. A DNA girase catalisa a super-helicoidização negativa no DNA de *E. coli*, e acredita-se que essa super-helicoidização negativa atrás das forquilhas de replicação guie o processo de desenrolamento porque a tensão super-helicoidal é reduzida pelo desenrolamento dos filamentos complementares.
- 10.25** A proteína DnaA inicia a formação da bolha de replicação por ligação às repetições de 9 pb de *OriC*. Sabe-se que a proteína DnaA é necessária para o processo de iniciação porque bactérias com mutações termossensíveis no gene do *dnaA* não iniciam a replicação de DNA em temperaturas restritivas.
- 10.27** Os nucleossomos e replissomos são grandes estruturas macromoleculares, e o empacotamento do DNA eucariótico em nucleossomos suscita a dúvida sobre o mecanismo pelo qual um replissomo pode ultrapassar o nucleossomo e replicar o DNA do nucleossomo durante esse processo. A solução mais óbvia para esse problema seria a desmontagem total ou parcial do nucleossomo para possibilitar a passagem do replissomo. Então, o nucleossomo seria remontado depois da passagem do replissomo. Um modelo popular é de desmontagem parcial do nucleossomo, possibilitando a passagem do replissomo (Figura 10.33B).
- 10.29** (1) A replicação de DNA geralmente é contínua nas células de procariotos, mas é restrita à fase S do ciclo celular em eucariotos. (2) A maioria dos cromossomos eucarióticos contém múltiplas origens de replicação, enquanto a maioria dos cromossomos procarióticos contém uma única origem de replicação. (3) Os procariotos usam dois complexos catalíticos que contêm a mesma

DNA polimerase para replicar o filamento líder e atrasado, ao passo que os eucariotos usam duas ou três DNA polimerases diferentes para a síntese dos filamentos líder e atrasado. (4) A replicação de cromossomos eucarióticos requer a desmontagem parcial e remontagem dos nucleossomos à medida que os replissomos se movem ao longo das moléculas de DNA parentais. Em procariotos, a replicação provavelmente inclui uma desmontagem/remontagem parcial semelhante de estruturas similares ao nucleossomo. (5) A maioria dos cromossomos procarióticos é circular e, portanto, não tem extremidades. A maioria dos cromossomos eucarióticos é linear e tem terminações especiais chamadas telômeros, acrescentadas às moléculas de DNA em replicação pela telomerase, uma enzima específica que contém RNA.

- 10.31** Os cromossomos de células haploides de levedura com a mutação *est1* tornam-se mais curtos a cada divisão celular. Por fim, a perda completa de telômeros causa instabilidade cromossômica e há morte celular em razão da deleção de genes essenciais perto das extremidades dos cromossomos.

## Capítulo 11

- 11.1** (a) O RNA contém o açúcar ribose, que tem um grupo hidroxila (OH) no carbono 2; o DNA contém o açúcar 2-desoxirribose, que tem apenas átomos de hidrogênio no carbono 2. O RNA geralmente tem a base uracila nas posições onde há timina no DNA. Algumas moléculas de DNA, porém, contêm uracila, e algumas de RNA contêm timina. Na maioria das vezes, o DNA é uma dupla hélice (molécula bifilamentar); o RNA é, com maior frequência, uma molécula unifilamentar; mas alguns DNA são unifilamentares e alguns RNA são bifilamentares. (b) A principal função do DNA é armazenar informações genéticas e transmitir essas informações de uma célula para outra e de uma geração para outra. O RNA armazena e transmite as informações genéticas em alguns vírus que não contêm DNA. Nas células que têm DNA e RNA: (1) o mRNA atua como intermediário na síntese de proteínas, levando as informações do DNA nos cromossomos para os ribossomos (locais de síntese de proteínas). (2) Os tRNA levam aminoácidos até os ribossomos e atuam no reconhecimento do códon durante a síntese de polipeptídios. (3) As moléculas de rRNA são componentes essenciais dos ribossomos. (4) Os snRNA são importantes componentes dos espliceossomos, e (5) os miRNA têm papéis essenciais na regulação da expressão gênica (Capítulo 19). (c) O DNA está localizado principalmente nos cromossomos (com uma parte nas organelas citoplasmáticas, como mito-



côndrias e cloroplastos), enquanto o RNA está em toda a célula.

### 11.3 3'-GACTA-5'.

11.5 A síntese proteica ocorre nos ribossomos. Nos eucariotos, a maior parte dos ribossomos está no citoplasma e fixada à extensa malha membranária de retículo endoplasmático. Também há alguma síntese proteica em organelas citoplasmáticas como cloroplastos e mitocôndrias.

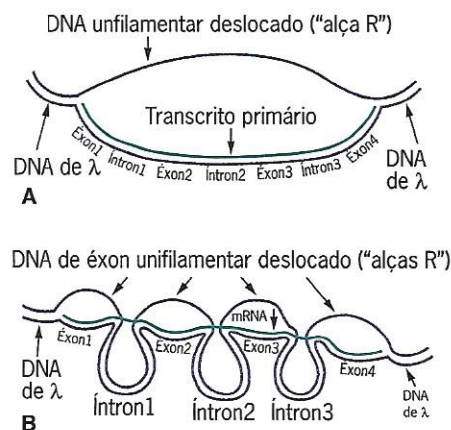
11.7 Tanto organismos procarióticos quanto eucarióticos têm RNA mensageiro, RNA transportador e RNA ribossômico. Além disso, os eucariotos contêm pequeno RNA nuclear (snRNA) e micro-RNA. As moléculas de RNA mensageiro levam informações genéticas dos cromossomos (onde são armazenadas) até os ribossomos no citoplasma (onde as informações são expressas durante a síntese proteica). A sequência linear de códons de trinucleotídeos em uma molécula de mRNA especifica a sequência linear de aminoácidos nos polipeptídeos produzidos durante a tradução desse mRNA. As moléculas de RNA transportador são pequenas (cerca de 80 nucleotídeos de comprimento) moléculas que transportam aminoácidos para os ribossomos e responsáveis pela especificidade de reconhecimento do códon durante a tradução. As moléculas de RNA ribossômico garantem parte da estrutura e função dos ribossomos; elas representam uma parte importante do aparelho necessário para a síntese de polipeptídeos. Pequenos RNA nucleares são componentes estruturais dos spliceossomos, que excisam sequências de íntrons não codificadores de transcritos de genes nucleares. Os micro-RNA participam da regulação da expressão gênica.

11.9 A "autorrecomposição" de precursores de RNA mostra que as moléculas de RNA também podem conter sítios catalíticos; essa propriedade não é restrita às proteínas.

11.11 Os íntrons de genes nucleares codificadores de proteínas de eucariotos superiores quase sempre começam (5') com GT e terminam (3') com AG. Além disso, a A subterminal 3' na "sequência TACTAAC" é totalmente conservada; essa A participa da formação de ligação durante a excisão do íntron.

11.13 (a) Sequência 5. Ela contém as sequências de íntrons conservadas: GU 5', AG 3' e sequência interna UACUAAAC que oferece um possível sítio de ligação para excisão do íntron. A sequência 4 tem GU 5' e AG 3', mas não tem A interna para o sítio de ligação durante a excisão do íntron. (b) 5'-UAGUCUCAA-3'; o suposto íntron de GU 5' a AG 3' foi removido.

### 11.15



11.17 Supondo-se que haja uma sequência -35 na região 5' em relação à sequência de consenso -10 da molécula de DNA, a sequência nucleotídica do transcrito será 5'-ACCCGACAUAAGCUACGAUGACGAUAAGC GACAUAAGC-3'.

11.19 Supondo-se que haja uma sequência CAAT na região 5' em relação à sequência TATA nesse segmento de DNA, a sequência nucleotídica do transcrito será 5'-ACCCGACAUAAGCUACGAUGACGAUA-3'.

11.21 Segundo o dogma central, as informações genéticas são armazenadas no DNA e transferidas do DNA para o RNA e do RNA para as proteínas durante a expressão gênica. Os vírus tumorais de RNA armazenam as informações genéticas em RNA, e essas informações são copiadas para o DNA pela enzima transcriptase reversa depois que o vírus infecta a célula hospedeira. Portanto, a descoberta dos vírus tumorais de RNA ou retrovírus - retro por causa da direção retrógrada do fluxo das informações genéticas - criou uma exceção ao dogma central.

11.23 A síntese de DNA, RNA e proteínas requer a síntese de longas cadeias de subunidades repetidas. Os três processos podem ser divididos em três estágios: iniciação da cadeia, alongamento da cadeia e finalização da cadeia.

11.25 Os transcritos primários de eucariotos sofrem maior processamento pós-transcrição que os de procariotos. Portanto, a maior diferença entre mRNA e transcritos primários ocorre em eucariotos. Em geral, o processamento do transcrito é restrito à excisão de sequências terminais nos procariotos. Já os transcritos eucarióticos geralmente são modificados por (1) excisão de sequências de íntron; (2) acréscimo de caps de 7-metilguanosina às terminações 5'; (3) acréscimo de caudas poli(A) às terminações 3'. Além disso, as sequências de alguns transcritos eucarióticos são modificadas por edição do RNA.

11.27 Em eucariotos, as informações genéticas são armazenadas no DNA nuclear, ao passo que as proteí-



nas são sintetizadas em ribossomos no citoplasma. Como os genes, que estão separados dos sítios de síntese proteica por uma membrana dupla – o envoltório nuclear – poderiam dirigir a síntese de polipeptídios sem algum tipo de intermediário que levasse do núcleo até o citoplasma as especificações para os polipeptídios? A princípio, os pesquisadores usaram precursores de RNA e proteínas marcados e autorradiografia para demonstrar que a síntese de RNA ocorria no núcleo e a síntese proteica, no citoplasma.

- 11.29** Um experimento simples de *pulse-labeling* (marcação por pulso) e *pulse-chase* (pulso e busca) demonstra que o RNA é sintetizado no núcleo e, depois, transportado para o citoplasma. Esse experimento é dividido em duas partes: (1) marcação de células eucarióticas por cultura em  $^3\text{H}$ -uridina durante alguns minutos e localização da radioatividade incorporada ou autorradiografia; (2) repetição do experimento, mas dessa vez com acréscimo de excesso de uridina não radioativa ao meio de cultura das células depois do período de marcação, aguardando o crescimento em meio não radioativo por aproximadamente uma hora. Depois, localização da radioatividade incorporada por meio da autorradiografia. A radioatividade estará localizada no núcleo quando as células em cultura forem marcadas com pulso de  $^3\text{H}$ -uridina e nos ribossomos no citoplasma no experimento de *pulse-chase*.
- 11.31** A primeira preparação de RNA polimerase provavelmente não tem a subunidade sigma e, por isso, inicia a síntese de cadeias de RNA em sítios aleatórios ao longo de ambos os filamentos do DNA de *argH*. A segunda preparação provavelmente contém a subunidade sigma e só inicia as cadeias de RNA no sítio usado *in vivo*, que é determinado pela posição das sequências -10 e -35 do promotor.
- 11.33** Sequências TATA e CAAT. As sequências TATA e CAAT geralmente estão centralizadas nas posições -30 e -80, respectivamente, em relação ao ponto de início (+1) da transcrição. A sequência TATA é responsável pelo posicionamento do ponto de início da transcrição; é o local de ligação do primeiro fator de transcrição basal que interage com o promotor. A sequência CAAT promove a eficiência da iniciação da transcrição.
- 11.35** Às vezes a edição de RNA leva à síntese de dois ou mais polipeptídios diferentes a partir de um único mRNA.
- 11.37** Esse zigoto provavelmente será inviável porque o produto gênico é essencial e é quase certo que a eliminação do sítio de recomposição 5' leve à síntese de produto gênico inativo.

## Capítulo 12

- 12.1** As proteínas são moléculas longas, semelhantes a cadeias, constituídas de aminoácidos unidos por ligações peptídicas. As proteínas são compostas de carbono, hidrogênio, nitrogênio, oxigênio e geralmente enxofre. Elas são responsáveis pela atividade enzimática e por grande parte da estrutura de organismos vivos. O DNA é constituído de fosfato, o açúcar pentose 2-desoxirribose e quatro bases orgânicas nitrogenadas (adenina, citosina, guanina e timina). O DNA armazena e transmite as informações genéticas na maioria dos organismos vivos. A síntese proteica tem interesse especial para os geneticistas porque as proteínas são os produtos gênicos primários – os intermediários essenciais por meio dos quais os genes controlam os fenótipos de organismos vivos.
- 12.3** Isso depende da definição de alelos. Se toda variação da sequência nucleotídica for considerada um alelo diferente, ainda que não haja modificação do produto gênico e do fenótipo do organismo mutante, o número de alelos estará diretamente relacionado com o tamanho do gene. Mas se for necessário que a modificação da sequência nucleotídica altere o produto gênico ou o fenótipo para que seja considerada um alelo distinto, haverá uma correlação positiva, mas não uma relação direta, entre o número de alelos de um gene e seu tamanho em pares de nucleotídios. A relação é mais provável em procariotos, nos quais a maioria dos genes não tem íntrons. Nos genes eucarióticos, as modificações da sequência nucleotídica nos íntrons geralmente são neutras; ou seja, não afetam a atividade do produto gênico nem o fenótipo do organismo. Portanto, no caso de genes eucarióticos com íntrons, pode não haver correlação entre o tamanho do gene e o número de alelos que modificam o fenótipo.
- 12.5** (a) Os códigos simples e duplo criam um máximo de 4 e  $(4)^2$  ou 16 códons, respectivamente. Portanto, nenhum código seria capaz de especificar todos os 20 aminoácidos. (b) 20. (c)  $(20)^{146}$ .
- 12.7** (a) O código genético é degenerado porque todos os 20 aminoácidos, exceto 2, são especificados por dois ou mais códons. Alguns aminoácidos são especificados por seis códons diferentes. A degeneração ocorre principalmente na terceira base ou base 3' dos códons. A "degeneração parcial" ocorre quando a terceira base do códon pode ser uma das duas purinas ou uma das duas pirimidinas e o códon ainda especifica o mesmo aminoácido. A "degeneração completa" ocorre quando a terceira base do códon pode ser qualquer uma das quatro bases e o códon ainda especifica o mesmo aminoácido. (b) O código é dito ordenado porque códons semelhantes (códons que têm apenas



uma base diferente) especificam aminoácidos quimicamente semelhantes. Por exemplo, os códons CUU, AUU e GUU especificam os aminoácidos de estrutura semelhante, leucina, isoleucina e valina, respectivamente. (c) O código parece ser quase totalmente universal. As exceções conhecidas à universalidade são linhagens com mutações supressoras que alteram a leitura de determinados códons (com baixa eficiência na maioria dos casos) e o uso de UGA como um códon de triptofano em mitocôndrias de leveduras e seres humanos.

- 12.9** His  $\rightarrow$  Arg é resultado de uma transição; His  $\rightarrow$  Pro exigiria uma transversão (não induzida por 5-bromouracila).
- 12.11** Os ribossomos têm diâmetro de 10 a 20 nm. Estão localizados principalmente no citoplasma celular. Nas bactérias, estão basicamente livres no citoplasma. Nos eucariotos, muitos ribossomos estão ligados ao retículo endoplasmático. Os ribossomos são estruturas complexas constituídas de mais de 50 polipeptídios diferentes e três a cinco moléculas diferentes de RNA.
- 12.13** As moléculas de RNA mensageiro levam informações genéticas dos cromossomos (onde são armazenadas) até os ribossomos no citoplasma (onde as informações são expressas durante a síntese proteica). A sequência linear de códons de trinucleotídeos em uma molécula de mRNA especifica a sequência linear de aminoácidos no(s) polipeptídio(s) produzido(s) durante a tradução desse mRNA. As moléculas de RNA transportador são pequenas (cerca de 80 nucleotídeos) moléculas que transportam aminoácidos para os ribossomos e responsáveis pela especificidade de reconhecimento de códon durante a tradução. As moléculas de RNA ribossômico garantem parte da estrutura e função dos ribossomos; elas representam uma parte importante do mecanismo necessário para a síntese de polipeptídios.
- 12.15** Uma aminoacil-tRNA sintetase específica catalisa a formação de um complexo aminoácido-AMP a partir do aminoácido apropriado e ATP (com liberação de pirofosfato). A mesma enzima então catalisa a formação do complexo aminoacil-tRNA, com liberação de AMP. As ligações aminoácido-AMP e aminoacil-tRNA são ligações fosfato de alta energia.
- 12.17** A hipótese da oscilação de Crick explica como o anticódon de determinado tRNA pode formar pares de bases com dois ou três códons de mRNA diferentes. Crick propôs que o pareamento de bases entre a base 5' do anticódon no tRNA e a base 3' do códon no mRNA era menos rigoroso que o normal e, portanto, possibilitava alguma "oscilação" nesse local. Logo, um único tRNA geralmente reconhece dois ou três dos códons semelhantes
- especificadores de determinado aminoácido (ver Tabela 12.2).
- 12.19** (a) Inosina. (b) Dois.
- 12.21** A tradução ocorre por mecanismos muito semelhantes em procariotos e eucariotos; no entanto, há algumas diferenças. (1) Em procariotos, a iniciação da tradução inclui pareamento de bases entre uma sequência conservada (AGGAGG) – a sequência de Shine-Dalgarno – no mRNA e uma sequência complementar próxima da extremidade 3' do rRNA 16S. Em eucariotos, o complexo de iniciação forma-se na extremidade 5' do transcrito quando uma proteína de ligação ao *cap* interage com a 7-metilguanosina no mRNA. Então, o complexo faz o exame processivo do mRNA e inicia a tradução (com algumas exceções) na sequência AUG mais próxima da terminação 5'. (2) Em procariotos, o grupo amino do iniciador metionil-tRNA<sup>Met</sup> é formilado; em eucariotos, o grupo amino de metionil-tRNA<sup>Met</sup> não é formilado. (3) Em procariotos, são necessários dois fatores de liberação (RF), que são proteínas solúveis, para a finalização da cadeia. RF-1 finaliza os polipeptídios em resposta aos códons UAA e UAG; RF-2 finaliza as cadeias em resposta aos códons UAA e UGA. Em eucariotos, um fator de liberação responde aos três códons de término.
- 12.23** Supondo-se 0,34 nm por par de nucleotídeos no B-DNA, um gene de 68 nm de comprimento conteria 200 pares de nucleotídeos. Em vista do código triplo, esse gene conteria  $200/3 = 66,7$  trinucleotídeos, e um deles especificaria obrigatoriamente a finalização da cadeia. Desprezando-se o trinucleotídeo parcial, esse gene codificaria no máximo 65 aminoácidos.
- 12.25**  $426 \text{ nucleotídeos} - 3 \times 141 = 423$  aminoácidos especificadores mais três (um códon) especificando a finalização da cadeia.
- 12.27** (a) Códons semelhantes frequentemente especificam aminoácidos iguais ou muito semelhantes. Assim, as substituições de apenas um par de bases frequentemente resultam na síntese de proteínas idênticas (degeneração) ou proteínas com substituições por aminoácidos muito semelhantes. (b) Leucina e valina têm estruturas e propriedades químicas muito semelhantes; ambas têm grupos laterais apolares e dobram-se basicamente nas mesmas estruturas tridimensionais quando presentes em polipeptídios. Portanto, substituições de valina por leucina ou de leucina por valina raramente alteram a função de uma proteína.
- 12.29** (a) Ribossomos e espliceossomos têm papéis essenciais na expressão gênica e são ambos estruturas macromoleculares complexas constituídas de moléculas de RNA e proteínas. (b) Os ribossomos



estão no citoplasma; os espliceossomos, no núcleo. Os ribossomos são maiores e mais complexos que os espliceossomos.

**12.31** Met-Ser-Ile-Cys-Leu-Phe-Gln-Ser-Leu-Ala-Ala-Gln-Asp-Arg-Pro-Gly.

**12.33** (UAG). Esse é o único códon sem sentido relacionado com os códons de triptofano, serina, tirosina, leucina, ácido glutâmico, glutamina e lisina pela substituição de apenas um par de bases em cada caso.

## Capítulo 13

**13.1** (a) Transição, (b) transição, (c) transversão, (d) transversão, (e) mudança de matriz de leitura, (f) transição.

**13.3** (a) Método *ClB*, (b) método ligado ao X (veja o Capítulo 6).

**13.5** Provavelmente não. O ser humano é maior que uma bactéria, tem mais células e vida mais longa. Se as frequências de mutação forem calculadas em termos de gerações celulares, as taxas em células humanas e bacterianas seriam semelhantes.

**13.7** O gene ligado ao X é transmitido pelas mães, e metade dos filhos de sexo masculino tem a doença. É difícil acompanhar essa doença em estudos de heredograma porque o gene é de natureza recessiva, a expressão tende a saltar gerações em uma linhagem familiar e os homens portadores do gene morrem. Uma explicação para a ocorrência esporádica da doença e a tendência de persistência do gene é que, por mutação, há acréscimo constante de novos genes defeituosos aos já existentes na população.

**13.9** O carneiro de pernas curtas pode ser cruzado com animais de pernas longas sem parentesco. Se a característica for expressa na primeira geração, pode-se presumir que é hereditária e determinada por um gene dominante. Por outro lado, se não aparecer na primeira geração, pode-se fazer o retrocruzamento dos carneiros da F1 com o genitor de pernas curtas. Se a característica for expressa em metade da prole do retrocruzamento, provavelmente a herança é recessiva simples. Se forem obtidos dois carneiros de pernas curtas de sexos diferentes, eles podem ser cruzados repetidas vezes para testar a hipótese de dominância. No caso em que a característica não for transmitida à prole desses cruzamentos, pode ser considerado ambiental ou dependente de algum mecanismo genético complexo que o teste simples usado nos experimentos não é capaz de identificar.

**13.11** Se genes mutacionais e antimutacionais operarem no mesmo sistema vivo, a seleção natural pode estabelecer uma taxa de mutação ideal

para um organismo específico em determinado ambiente.

**13.13** (a) Sim. (b) O bloqueio provocaria o acúmulo de fenilalanina e a diminuição da quantidade de tirosina, um resultado que seria esperado em várias expressões fenotípicas diferentes.

**13.15**

| Aminoácido      | mRNA  | DNA                                 |
|-----------------|-------|-------------------------------------|
| Ácido glutâmico | —GAA→ | —GAA→<br>←CTT← Filamento transcrito |
| Valina          | —GUA→ | —GTA→<br>←CAT←<br>Mutação           |
| Lisina          | —AAA→ | —AAA→<br>←TTT←                      |

**13.17** Mutações: transições, transversões e mudanças de matriz de leitura.

**13.19** 3%; 4%; 6%.

**13.21** O iodo radioativo é concentrado nos organismos vivos e nas cadeias alimentares.

**13.23** ( $x^+ m^+ z$ ) ( $x^+ m^+ z^+$ ) ( $x m^+ z$ ) ( $x m z^+$ ) ou equivalente.

**13.25** Transições.

**13.27** O ácido nitroso atua como mutágeno, esteja o DNA em replicação ou não, e produz transições de A para G ou de C para T, ao passo que a 5-bromouracila não afeta o DNA em replicação, mas atua durante a replicação, causando transições GC ↔ AT. É preciso que a 5-bromouracila seja incorporada ao DNA durante o processo de replicação para induzir erro de pareamento de bases e, portanto, mutações.

**13.29** A 5-BU causa transições GC ↔ AT. Portanto, a 5-BU pode reverter quase todas as mutações que induz por meio de estímulo do processo de transição, que é o inverso do que produziu a mutação. Já as mutações espontâneas abrangem a transversão, a mudança de matriz de leitura, a deleção e outros tipos de mutação, inclusive a transição. Apenas as transições espontâneas têm reversão aumentada depois do tratamento com 5-BU.

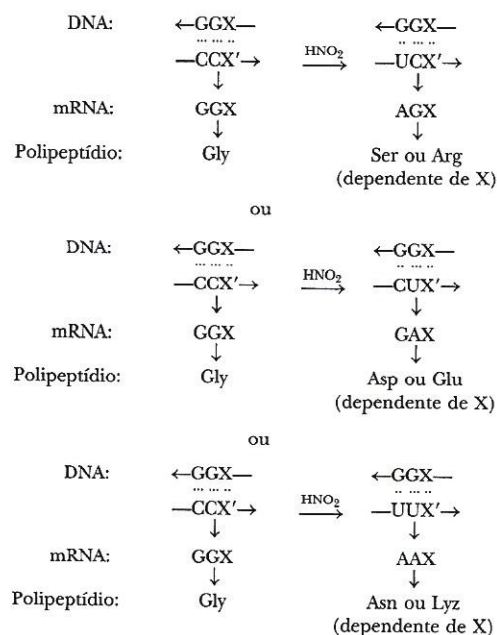
**13.31** (a) Mudança de matriz de leitura por inserção de C no 9º, 10º e 11º nucleotídeos a partir da extremidade 5'. (b) Normal: 5'-AUGCCGUACUGCCAGCUAACUGCUAAAGAACAAUUA-3'. Mutante: 5'-AUGCCCGUACUGCCAGCUAACUGCUAAAGAACAAUUA-3'. (c) Normal: NH<sub>2</sub>-Met-Pro-Tyr-Cys-Gln-Leu-Thr-Ala-Lys-Glu-Gln-Leu. Mutante: NH<sub>2</sub>-Met-Pro-Val-Leu-Pro-Ala-Asn-Cys.

**13.33** Não. A substituição leucina → prolina seria mais frequente. A substituição Leu (CUA) → 5-BU → Pro (CCA) ocorre por transição de um único par de bases, enquanto Leu (CUA) → 5-BU Ser (UCA) requer transição de dois pares de bases. Lembre-se



de que a 5-bromouracila (5-BU) induz apenas transições (ver Figura 13.7).

### 13.35 Sim:



*Nota:* O X na terceira posição em cada códon de mRNA e em cada trinca de pares de bases no DNA indica que há degeneração completa na terceira base do códon de glicina. Qualquer que seja a base presente no códon ele ainda especificará a glicina.

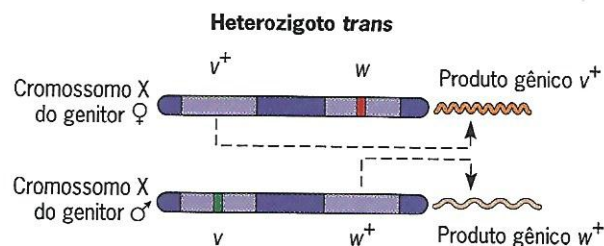
13.37 Substituições Tyr  $\rightarrow$  Cys; a substituição de Tyr por Cys requer uma transição, que é induzida pelo ácido nitroso. A substituição de Tyr por Ser exigiria uma transversão, e o ácido nitroso não induz transversão.

13.39 5'-UGG-UGG-UGG-AUG-CGA ou AGA-GAA ou GAG-UGG-AUG-3'.

13.41 Dois genes; as mutações 1, 2, 3, 4, 5, 6 e 8 estão em um gene; a mutação 7 está em outro gene.

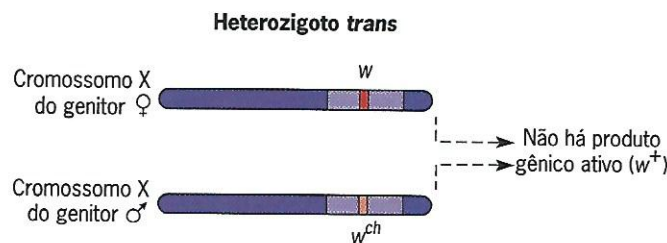
13.43 O teste de complementação para alelismo requer que se insiram pares de mutações em um protoplasma comum na configuração *trans* e se determine se os heterozigotos *trans* produzidos têm fenótipo selvagem ou mutante. Se estiverem em genes diferentes, as duas mutações se complementarão, porque as cópias selvagens de cada gene especificam produtos gênicos ativos (ver Figura 13.23A). Mas se as duas mutações estiverem no mesmo gene, as duas cópias do gene no heterozigoto *trans* especificarão produtos gênicos anômalos, resultando em fenótipo mutante (ver Figura 13.23B). Quando há complementação, o heterozigoto *trans* tem fenótipo selvagem. O teste de complementação ou *trans* pode ser usado para verificar se duas mutações recessivas estão localizadas no mesmo gene ou em dois genes diferentes.

Como as mutações de interesse estão ligadas ao sexo, toda a prole masculina terá fenótipo igual ao da mãe. Eles são hemizigotos, com um cromossomo X herdado da mãe. Já a prole feminina é constituída de heterozigotos *trans*. A prole feminina do cruzamento de fêmeas de olhos brancos e machos de olhos vermelhão tem olhos vermelhos, o fenótipo selvagem. Portanto, as mutações *white* e *vermilion* estão em genes diferentes, como ilustra o diagrama a seguir:



A complementação produz fenótipo selvagem; os produtos gênicos  $v^+$  e  $w^+$  são produzidos no heterozigoto *trans*.

A prole feminina do cruzamento de uma fêmea de olhos brancos e um macho de olhos branco-cereja tem olhos cereja-claros (fenótipo mutante), e não olhos vermelhos de tipo selvagem como no primeiro cruzamento. Como o heterozigoto *trans* tem fenótipo mutante, duas mutações, *white* e *white cherry*, estão no mesmo gene.



Não há produto gênico  $w^+$ ; portanto, o fenótipo é mutante.

## Capítulo 14

14.1 (a) Ambas introduzem nova variabilidade genética na célula. Nos dois casos, apenas um gene ou um pequeno segmento de DNA representando uma pequena fração do genoma total é modificado ou acrescentado ao genoma. A grande maioria dos genes do organismo continua idêntica. (b) É mais provável que a introdução de moléculas de DNA recombinantes, se originadas de uma espécie muito diferente, resulte em um novo produto gênico funcional na célula, se for possível a expressão do gene (ou genes) introduzido no protoplasma estranho. A introdução de moléculas de DNA recombinantes é mais semelhante à mutação por duplicação (veja o Capítulo 6) que aos outros tipos de mutações.

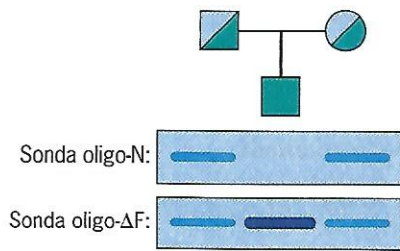
14.3 (a)  $(1/4)^4 = 1/256$ ; (b)  $(1/4)^6 = 1/4.096$ .



- 14.5 As técnicas de recombinação do DNA e clonagem gênica possibilitam que os geneticistas isolem praticamente qualquer gene ou sequência de DNA de interesse e procedam à sua caracterização estrutural e funcional. É possível obter grande quantidade de determinado gene na forma pura, o que torna possível identificar sua sequência de pares de nucleotídeos ("sequenciar o gene" no jargão de laboratório). A partir da sequência nucleotídica de um gene e do nosso conhecimento do código genético, os geneticistas conseguem prever a sequência de aminoácidos do polipeptídeo codificada pelo gene. Com o auxílio de um subclone apropriado do gene como sonda de hibridização em análises por *Northern blot*, os geneticistas são capazes de identificar os tecidos em que o gene é expresso. Com base na sequência de aminoácidos prevista de um polipeptídeo codificado por um gene, os geneticistas podem sintetizar oligopeptídeos e usá-los para produzir anticorpos que, por sua vez, podem ser usados para identificar o verdadeiro produto do gene e localizá-lo em células ou tecidos do organismo. Assim, as tecnologias de recombinação do DNA e clonagem gênica são métodos muito eficientes para o estudo do controle genético de quase todos os processos biológicos. Esses métodos tiveram papéis importantes no explosivo progresso ocorrido no campo da biologia durante as três últimas décadas.
- 14.7 Acredita-se que as endonucleases de restrição atuem como um tipo de sistema imune primitivo nos microrganismos que as produzem – protegendo seu material genético contra a "invasão" por DNA estranho de vírus ou outros patógenos ou por DNA do ambiente que poderia ser absorvido pelo microrganismo. É claro que esses microrganismos não têm um sistema imune sofisticado como os dos animais superiores (Capítulo 20).
- 14.9 Um DNA estranho clonado com auxílio de uma enzima que produz extremidades complementares monofilamentares sempre pode ser excisado do vetor de clonagem por clivagem pela mesma enzima de restrição originalmente usada para cloná-lo. Se um fragmento de *HindIII* contendo seu gene favorito foi clonado em DNA vetor Bluescript clivado por *HindIII*, será ladeado no clone Bluescript recombinante por sítios de clivagem para *HindIII*. Portanto, pode-se excisar esse fragmento de *HindIII* por digestão do clone Bluescript com endonuclease *HindIII*.
- 14.11 A maioria dos genes de vegetais e animais superiores contém sequências de íntrons não codificadoras. Essas sequências de íntrons estão presentes nos clones genômicos, mas não em clones de cDNA, porque os cDNA são sintetizados a partir de moldes de mRNA e as sequências de íntrons são removidas durante o processamento dos transcritos primários para produzir mRNA maduros.
- 14.13 O gene *gln2* do milho contém muitos íntrons, e um dos íntrons contém um sítio de clivagem para *HindIII*. As sequências de íntrons (e, portanto, o sítio de clivagem para *HindIII*) estão ausentes em sequências de mRNA e, portanto, também estão ausentes em clones de sequência completa de cDNA de *gln2*.
- 14.15 (a) Os métodos de *Southern*, *Northern* e *Western blot* têm uma etapa em comum, a transferência de macromoléculas (DNA, RNA e proteínas, respectivamente) separadas por eletroforese em gel para um suporte sólido – geralmente uma membrana de nitrocelulose ou náilon – para análise complementar. (b) A principal diferença entre essas técnicas é a classe de macromolécula separada durante a etapa de eletroforese: DNA em *Southern blots*, RNA em *Northern blots* e proteína em *Western blots*.
- 14.17 Todos os vetores de clonagem modernos contêm um "sítio de *polylinker*" ou "sítio de clonagem múltipla" (MCS) – um grupo de sítios de clivagem simples para várias diferentes endonucleases de restrição em uma região não essencial do vetor na qual se pode inserir o DNA estranho. Em geral, quanto maior é a complexidade do MCS – ou seja, quanto mais sítios de clivagem para endonuclease de restrição estão presentes – maior é a utilidade do vetor para clonagem de uma grande variedade de diferentes fragmentos de restrição. Veja, por exemplo, o MCS presente no plasmídeo Bluescript II mostrado na Figura 14.3.
- 14.19 Como se conhecem as sequências de pares de nucleotídeos tanto do gene *CF* normal quanto do gene *CF*  $\Delta 508$  mutante, é possível sintetizar oligonucleotídeos marcados e usá-los como sondas de hibridização para detectar a presença de cada alelo (normal e  $\Delta 508$ ). Em condições de hibridização muito rigorosas, cada sonda hibridiza-se somente com o alelo de *CF* com o qual tem perfeita complementaridade. Como as sequências do gene *CF* que ladeiam o sítio  $\Delta 508$  são conhecidas, é possível sintetizar iniciadores de PCR oligonucleotídicos e usá-los para amplificar por PCR esse segmento do DNA obtido a partir de pequenos explantes teciduais de supostos pacientes com FC e seus parentes. Em seguida, o DNA amplificado pode ser separado por eletroforese em gel de agarose, transferido para membranas de náilon e hibridizado com as respectivas sondas oligonucleotídicas marcadas, com detecção de cada alelo de *CF* presente por autorradiografia. Veja uma demonstração da utilidade desse procedimento em Em foco: Detecção de um gene mutante causador de fibrose cística. No procedimento descrito, empregaram-se duas sondas oligonucleotídicas sintéticas – oligo-N = 3'-CTTTTATAGTAGAAACCAC-5' e oligo- $\Delta F$  = 3'-TTCTTTTATAGTA-ACCACAA-5' (o travessão indica os nucleotídeos deletados no alelo



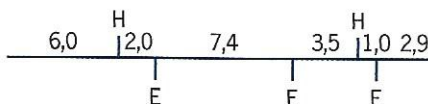
mutante *CF*Δ508) – para analisar o DNA de pacientes com FC e seus genitores. Nas famílias com FC confirmada, os resultados dessas hibridizações por *Southern blot* com as sondas oligo-N (normal) e oligo-ΔF (*CF*Δ508) marcadas geralmente foram:



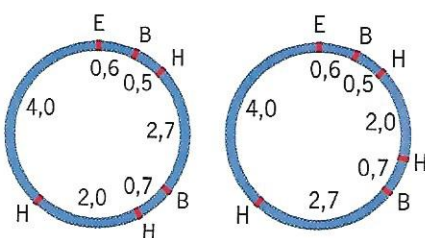
Pai e mãe eram heterozigotos para o alelo normal de *CF* e o alelo *CF*Δ508 mutante, como é esperado em um traço recessivo raro, e o paciente com FC era homozigoto para o alelo *CF*Δ508. Nessas famílias, a expectativa é de que um quarto das crianças seja homozigoto para o alelo mutante Δ508 e tenha sintomas de FC, enquanto três quartos sejam normais (não tenham FC). No entanto, a expectativa é de que dois terços dessas crianças normais sejam heterozigotos e transmitam o alelo para seus filhos. Apenas um quarto das crianças dessa família seria homozigoto para o alelo normal de *CF* e não correria o risco de transmitir o gene *CF* mutante para a prole. Observe que o método de rastreamento descrito aqui pode ser usado para determinar quais das crianças normais são portadoras do alelo *CF*Δ508: ou seja, o gene mutante pode ser detectado em heterozigotos e homozigotos.

- 14.21** A seleção genética é a técnica mais eficiente de clonagem de genes desse tipo. Prepare uma biblioteca genômica em um vetor de expressão como o Bluescript (ver Figura 14.3) usando DNA da linhagem de *Shigella dysenteriae* resistente à canamicina. Então, pesquise o gene de resistência à canamicina na biblioteca transformando células de *E. coli* sensíveis à canamicina com os clones na biblioteca e plaqueando as células transformadas em meio contendo canamicina. Somente as células transformadas pelo gene de resistência à canamicina produzem colônias na presença de canamicina.

**14.23**

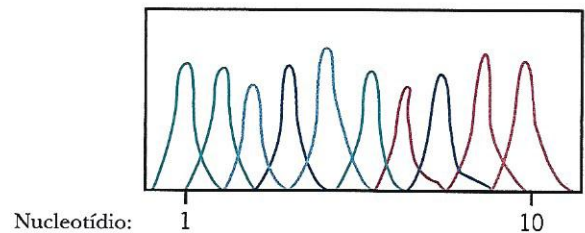


- 14.25** Há dois mapas de restrição possíveis para esses dados, mostrados adiante:



Os sítios de clivagem para as enzimas de restrição *Bam*HI, *Eco*RI e *Hind*III são indicados pelas letras B, E e H, respectivamente. Os números indicam as distâncias em quilopares de bases.

**14.27**



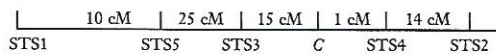
## Capítulo 15

- 15.1** As distâncias de mapa genético são determinadas por frequências de *crossing over*. Os mapas citogenéticos são fundamentados na morfologia do cromossomo ou nas características físicas dos cromossomos. Os mapas físicos são baseados em distâncias físicas reais – o número de pares de nucleotídeos (0,34 nm por pb) – que separam os marcadores genéticos. Se um gene ou outra sequência de DNA de interesse está perto de um gene mutante, uma banda específica em um cromossomo ou um fragmento de restrição de DNA específico, pode-se usar esse marcador genético ou físico (mutação, banda ou fragmento de restrição) para iniciar uma caminhada no cromossomo até o gene de interesse (Figura 15.7).

- 15.3** Contig (clones *contíguos*) é um mapa físico de um cromossomo ou de parte de um cromossomo preparado a partir de um conjunto de clones de DNA genômicos coincidentes. RFLP (polimorfismo de comprimento de fragmento de restrição) é uma variação no comprimento de um fragmento de restrição específico excisado de um cromossomo por digestão por uma ou mais endonucleases de restrição. VNTR (repetição em série em número variável) é uma sequência curta de DNA presente no genoma como repetições consecutivas e em número muito variável. STS (sítio marcado por sequência) é uma sequência exclusiva de DNA mapeada em um sítio específico de um cromossomo. EST (etiqueta de sequência expressa) é uma sequência de cDNA – uma sequência genômica que é transcrita. Os mapas de contig possibilitam que pesquisadores obtenham clones com genes de interesse diretamente dos centros de estoque de DNA – “clonagem por telefone”. Os RFLP são usados para construir os mapas genéticos de alta densidade necessários para clonagem posicional. As VNTR são RFLP particularmente úteis empregadas na identificação de múltiplos sítios em genomas. Os STS e as EST são sondas moleculares que podem ser usadas para iniciar caminhadas no cromossomo até genes vizinhos de interesse.



## 15.5 (a)



(b)  $3,3 \times 10^9 \text{ pb} / 3,3 \times 10^3 \text{ cM} = 1 \times 10^6 \text{ pb/cM}$ . O comprimento de mapa total é 65 cM, que é igual a aproximadamente  $65 \times 10^6$  ou 65 milhões de pb.

(c) O gene do câncer (C) e STS4 são separados por 1 cM ou cerca de um milhão de pares de bases.

15.7 Em um clone do gene disponível, a hibridização *in situ* com fluorescência (FISH) pode ser usada para identificar o cromossomo humano que tem o gene e para localizar o gene no cromossomo. As cópias unifilamentares dos clones são acopladas a uma sonda fluorescente e hibridizadas com o DNA desnaturado em cromossomos dispersos sobre uma lâmina. Depois da hibridização, a sonda livre é removida por lavagem, e a localização da sonda fluorescente é determinada por fotografia usando microscopia de fluorescência (veja o Apêndice C: Hibridização *in situ*).

15.9 As repetições em série em número variável (VNTR) são constituídas de sequências repetidas com 10 a 80 pares de nucleotídeos, e as repetições curtas em série (STR) são constituídas de sequências repetidas com 2 a 10 pares de nucleotídeos.

15.11 A resolução do mapeamento de híbridos por radiação é maior que a do mapeamento de híbrido de células somáticas tradicional porque a frequência de recombinação é muito aumentada em híbridos por radiação pelo uso de raios X para fragmentar os cromossomos humanos antes da fusão celular. O fundamento do mapeamento de híbridos por radiação é que a probabilidade de quebrar uma molécula de DNA na região entre os dois genes e, assim, separá-los é diretamente proporcional à distância física (número de pares de bases) entre eles.

15.13 Os objetivos do Projeto Genoma Humano eram preparar mapas genéticos e físicos que mostrassem as localizações de todos os genes no genoma humano e determinar as sequências nucleotídicas dos 24 cromossomos no genoma humano. Esses mapas e as sequências nucleotídicas dos cromossomos humanos ajudaram os cientistas a identificar genes mutantes causadores de doenças hereditárias. Espera-se que a identificação desses genes mutantes causadores de doença leve ao tratamento eficaz, entre eles a terapia gênica, de pelo menos algumas dessas doenças no futuro. Os possíveis usos indevidos desses dados incluem invasão de privacidade pelo governo e por empresas – sobretudo agências de emprego e seguradoras. Não se podem negar oportunidades educacionais, emprego ou seguro às pessoas em razão de doen-

ças hereditárias ou mutações gênicas que predisponham a anormalidades mentais ou físicas.

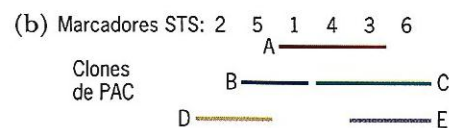
15.15 É mais provável que haja uma EST do que um RFLP em gene humano causador de doença. As EST correspondem às sequências expressas em um genoma. Os RFLP ocorrem em todo o genoma, tanto em sequências expressas quanto não expressas. Como menos de 2% do genoma humano codifica proteínas, a maioria dos RFLP ocorre no DNA não codificador.

15.17 (a) Segmento 5; (b) segmento 4; (c) segmento 1, 6 ou 10.

15.19 A principal vantagem dos *chips* gênicos como instrumento de hibridização por microarranjo é que um único *chip* gênico pode ser usado para a quantificação simultânea de milhares de sequências nucleotídicas. A tecnologia do *chip* gênico possibilita que os pesquisadores investiguem os níveis de expressão de grande quantidade de genes com maior eficiência do que era possível usando procedimentos de microarranjo anteriores.

15.21 As sequências de DNA em bibliotecas de cDNA específicas de cromossomo humano podem ser acopladas a corantes fluorescentes e hibridizadas *in situ* com os cromossomos de outros primatas. Os padrões de hibridização podem ser usados para detectar alterações na estrutura genômica ocorridas durante a evolução das várias espécies de primatas a partir de ancestrais comuns. Essas comparações são particularmente eficazes na detecção de novas relações de ligação decorrentes de translocações e fusões cêtricas.

15.23 (a) Ordem dos sítios STS: 2-5-1-4-3-6.



15.25 Todas as sequências identificadas pela busca megablast codificam proteínas histonas H2a. A sequência consultada é idêntica à sequência codificadora do gene *H2aV* de *Drosophila melanogaster* (membro da família de genes codificadores das proteínas histonas H2a). A sequência consultada codifica um polipeptídeo histona H2a de *Drosophila* designado variante V. As mesmas sequências do banco de dados são identificadas quando se usa metade ou um quarto da sequência nucleotídica dada na busca megablast. Sequências de apenas 15 a 20 nucleotídeos podem ser usadas para identificar o gene de *Drosophila* codificador da variante de histona H2a. No entanto, os resultados variam dependendo da sequência nucleotídica específica usada na consulta.



- 15.27 A matriz de leitura 5' → 3' número 1 tem uma grande matriz aberta de leitura com um códon de metionina perto da extremidade 5'. Pode-se verificar que essa é a matriz de leitura correta pelo uso do produto da tradução previsto para consulta em um dos bancos de dados de proteínas (veja a Questão 15.26).

## Capítulo 16

- 16.1 As ilhas de CpG são aglomerados de citosinas e guaninas, com frequência localizadas em direção 5' logo ao lado das regiões codificadoras de genes humanos. Sua presença em sequências nucleotídicas pode oferecer pistas sobre a localização de genes em cromossomos humanos.
- 16.3 O gene *CF* foi identificado por clonagem com base na posição no mapa, e as sequências nucleotídicas de cDNA de *CF* foram usadas para prever a sequência de aminoácidos do produto do gene *CF*. Uma busca em computador nos bancos de dados de proteínas mostrou que o produto do gene *CF* era semelhante a várias proteínas de canal iônico. Esse resultado voltou a atenção dos cientistas estudiosos da fibrose cística para as proteínas participantes do transporte de sais entre as células e levou à descoberta de que o produto do gene *CF* era um regulador da condutância transmembrana – agora denominado proteína CFTR.
- 16.5 Iniciadores oligonucleotídicos complementares às sequências de DNA nos dois lados (5' e 3') da região da repetição CAG no gene *MD* podem ser sintetizados e usados para amplificar a região de repetição por PCR. Um iniciador tem de ser complementar a uma região 5' do filamento molde, e o outro iniciador tem de ser complementar a uma região 3' do filamento não molde. Após amplificação, podem-se determinar os tamanhos das regiões de repetição CAG por eletroforese em gel (Figura 16.2). A medida do comprimento da repetição trinucleotídica pode ser feita pela inclusão no gel de regiões de repetições cujos comprimentos são conhecidos. Se houver menos de 30 cópias da repetição trinucleotídica em cada cromossomo, o recém-nascido, feto ou pré-embrião é homozigoto para um alelo *MD* tipo selvagem ou heterozigoto para dois alelos *MD* tipo selvagem diferentes. Se houver mais de 50 cópias da repetição em cada cromossomo homólogo, o indivíduo, feto ou célula é homozigoto para um alelo *MD* mutante dominante ou heterozigoto para dois alelos mutantes diferentes. Se um cromossomo tiver menos de 30 cópias da repetição CAG e o cromossomo homólogo tiver mais de 50 cópias, o recém-nascido, feto ou pré-embrião é heterozigoto, com um alelo *MD* selvagem e um alelo *MD* mutante.
- 16.7 Os sinais de início e término da transcrição e de início da tradução em eucariotos e em procariotos como a *E. coli* são diferentes. Portanto, para produzir uma proteína humana em *E. coli*, é preciso juntar a sequência codificadora do gene humano a sinais reguladores apropriados de *E. coli* – sequências promotora, finalizadora da transcrição e iniciadora da tradução. Além disso, se o gene contiver íntrons, é preciso removê-los ou usar a sequência codificadora de cDNA, porque a *E. coli* não tem os espliceossomos necessários para a excisão de íntrons dos transcritos de genes nucleares. Além disso, muitas proteínas eucarióticas passam por eventos de processamento após a tradução que não ocorrem nas células procarióticas. Essas proteínas são produzidas com mais facilidade em células eucarióticas transgênicas em cultura.
- 16.9 Onze, com variação, em múltiplos de três, de 15 a 45 nucleotídeos de comprimento.
- 16.11 Os perfis de DNA são os padrões específicos (1) de picos presentes em eletroferogramas de STR ou VNTR cromossômicos amplificados por PCR usando iniciadores marcados com corantes fluorescentes e separados por eletroforese capilar em gel (ver Figuras 16.11 e 16.12) ou (2) de bandas em *Southern blots* de DNA genômico digerido por enzimas de restrição específicas e hibridizado com sequências STR ou VNTR apropriadas (Figura 16.10). Os perfis de DNA, assim como as impressões digitais epidérmicas, são usados como provas de identidade ou de exclusão de identidade em processos judiciais. Os geneticistas manifestaram suas preocupações sobre os usos estatísticos dos dados do perfil de DNA. Eles questionaram, em especial, alguns métodos usados para calcular a probabilidade de que o DNA de outra pessoa, além do suspeito, pudesse ter produzido o perfil observado. Essas preocupações fundamentam-se, em parte, na ausência de bancos de dados adequados para várias subpopulações humanas e na ausência de informações precisas sobre o grau de variabilidade de perfis de DNA em indivíduos de diferentes etnias. Uma maneira de enfrentar esse problema foi a aquisição de dados sobre as frequências de perfis em diferentes populações e grupos étnicos do mundo todo.
- 16.13 A contaminação das amostras de sangue introduziria maior variabilidade nos perfis de DNA. A consequência seria a ausência de correspondência alélica dos perfis obtidos das amostras de sangue e do réu. Os erros por mistura causariam a absolvição de um culpado, mas não a condenação de um inocente. Somente a identificação errada das amostras poderia incriminar uma pessoa inocente.
- 16.15 A sondagem de *Southern blots* de DNA digerido por enzima de restrição dos vegetais transgênicos com transgene marcado com <sup>32</sup>P pode comprovar as múltiplas inserções, mas não revelaria a localização genômica dos insertos. A hibridização *in situ*



com fluorescência (FISH) é um procedimento eficiente para determinar a localização genômica dos insertos gênicos. A FISH é usada para observar a localização de transgenes em cromossomos (veja o Apêndice C).

- 16.17** Camundongos transgênicos geralmente são produzidos pela microinjeção dos genes de interesse em pronúcleos de ovos fertilizados ou pela infecção de embriões pré-implantação por vetores retrovirais contendo os genes de interesse. Os camundongos transgênicos são instrumentos muito úteis para estudar a expressão gênica, o desenvolvimento dos mamíferos e o sistema imune dos mamíferos. Os camundongos transgênicos são muito importantes na medicina; eles constituem o sistema-modelo mais estreitamente relacionado com os seres humanos. Eles foram e, sem dúvida, continuarão sendo valiosos no desenvolvimento dos métodos e da tecnologia que serão usados em terapia gênica humana no futuro.
- 16.19** As proteínas modificadas após a tradução podem ser produzidas em células eucarióticas transgênicas em cultura ou em vegetais e animais transgênicos. Na verdade, produziram-se ovelhas transgênicas que secretam no leite o fator da coagulação sanguínea IX e  $\alpha$ 1-antitripsina humanos. Essas ovelhas foram produzidas pela fusão das sequências codificadoras dos respectivos genes com uma sequência de DNA que codifica o peptídeo sinalizador necessário para secreção, seguida por introdução do gene quimérico em ovócitos fertilizados que foram implantados e deram origem a animais transgênicos. Em princípio, essa técnica poderia ser usada para produzir qualquer proteína de interesse.
- 16.21** O vetor descrito contém o gene *HGH*, mas não contém um promotor de *HGH* de mamíferos que controle a expressão do transgene nos tecidos apropriados. A construção de vetores que contenham uma sequência promotora de *HGH* de mamífero em posição correta deve resultar em camundongos transgênicos nos quais a síntese de *HGH* seja restrita à hipófise.
- 16.23** A RNAi é o uso de RNA bifilamentar, no qual um filamento é complementar ao mRNA e o outro filamento é equivalente ao mRNA, para silenciar a expressão de genes específicos. A RNAi emprega o complexo de silenciamento induzido por RNA (RISC) para bloquear a expressão gênica (Figura 16.23).
- 16.25** Os vegetais têm uma vantagem em relação aos animais porque, uma vez induzidas, as mutações por inserção podem ser armazenadas por longos períodos e distribuídas para os pesquisadores como sementes latentes.
- 16.27** (a) Primeiro seria preciso consultar o *site* do Salk Institute's Genome Analysis Laboratory para verificar se já foi identificada a inserção de um T-DNA

ou transpôson nesse gene (veja o Problema 16.28). Em caso afirmativo, basta solicitar sementes da linhagem transgênica ao *Arabidopsis* Biological Resource Center da Ohio State University. Se não houver inserção disponível no gene, pode-se verificar onde está mapeada no genoma e usar transpôsons que saltam preferencialmente para sítios vizinhos para identificar uma nova mutação por inserção (veja <http://www.arabidopsis.org/abrc/ima.jsp>). (b) Pode-se construir um gene que tenha sequências *sense* e *antisense* transcritas em uma única molécula de mRNA (Figura 16.23B), introduzi-lo em vegetais *Arabidopsis* por transformação mediada por *A. tumefaciens* e estudar seu(s) efeito(s) sobre a expressão do gene e o fenótipo de vegetais transgênicos. O transcrito formará um grampo com pareamento parcial de bases que entrará na via de silenciamento RISC e bloqueará a expressão do gene (Figura 16.23B).

## Capítulo 17

- 17.1** O par na opção (d) é de repetições invertidas e, portanto, poderia ser qualificado.
- 17.3** A resistência ao segundo antibiótico foi adquirida por transferência gênica conjugativa entre os dois tipos de células.
- 17.5** Na primeira cepa, o fator F integrou-se ao cromossomo por recombinação com o elemento IS entre os genes *C* e *D*. Na segunda cepa, integrou-se por recombinação com o elemento IS entre os genes *D* e *E*. As duas cepas transferem seus genes em ordens diferentes porque os dois elementos IS cromossômicos estão em sentidos opostos.
- 17.7** Não. *IS1* e *IS2* são mobilizados por diferentes transposases.
- 17.9** Mutação *tnpA*: não; mutação *tnpR*: sim.
- 17.11** Muitos transpôsons bacterianos têm genes de resistência a antibióticos, e é relativamente simples o deslocamento desses genes de uma molécula de DNA para outra. As moléculas de DNA que adquirem genes de resistência podem ser passadas para outras células em uma população de bactérias, tanto de modo vertical (por descendência) quanto horizontal (por transferência conjugativa). Com o passar do tempo, a exposição contínua a um antibiótico seleciona as células que adquiriram um gene para resistência a esse antibiótico. Portanto, o antibiótico não será mais útil no combate a essas bactérias.
- 17.13** A mutação *c<sup>m</sup>* é causada por uma inserção de *Ds* ou de *Ac*.
- 17.15** O alelo *Bz* de herança paterna foi inativado por inserção de um elemento transponível.
- 17.17** Cruzamento de machos disgênicos (extremamente mutáveis) com um cromossomo X selvagem e fêmeas homozigotas para um cromossomo X



balanceador; depois, cruzamento individual das filhas heterozigotas da  $F_1$  com seus irmãos e pesquisa de machos da  $F_2$  que não tenham o cromossomo balanceador para fenótipos mutantes, inclusive incapacidade de sobreviver (letalidade). As mutações identificadas nessa pesquisa provavelmente são causadas por inserções de elemento  $P$  em genes ligados ao  $X$ .

- 17.19 A transposição requer fatores produzidos pelo genoma da mosca; outros insetos aparentemente são incapazes de produzir esses fatores.
- 17.21 Por *crossing over* entre os LTR de um elemento  $Ty1$ .
- 17.23 Hibridização *in situ* com cromossomos politênicos usando uma sonda *TART* (veja o Apêndice C).
- 17.25 *TART* e *HeTA* restauram as extremidades dos cromossomos de *Drosophila*.
- 17.27 O elemento *Sleeping Beauty* poderia ser usado como vetor de transformação em vertebrados de modo muito semelhante ao uso do elemento  $P$  em *Drosophila*. O gene *gfp* poderia ser inserido entre as extremidades do elemento *Sleeping Beauty* e injetado em ovócitos ou embriões junto com um elemento *Sleeping Beauty* intacto capaz de codificar a transposase do elemento. Se a transposase produzida no ovócito ou embrião injetado atuar sobre o elemento contendo o gene *gfp*, pode causar a inserção desse gene no DNA genômico. Então, se o ovócito ou embrião der origem a um adulto, pode-se cruzá-lo para verificar se há transmissão de um transgene *Sleeping Beauty/gfp* para a geração seguinte. Desse modo, seria possível obter cepas de camundongos ou peixes-zebras que expressassem o gene *gfp*.

## Capítulo 18

- 18.1 Por meio do estudo da síntese ou ausência de síntese da enzima em células cultivadas em meios de composição química definida. Se a enzima for sintetizada apenas na presença de determinado metabólito ou de um grupo específico de metabólitos, provavelmente é induzível. Se for sintetizada na ausência, mas não na presença de determinado metabólito ou grupo de metabólitos, provavelmente é repressível.

### 18.3

| Gene ou elemento regulador | Função                                                    |
|----------------------------|-----------------------------------------------------------|
| (a) Gene regulador         | Codifica o repressor                                      |
| (b) Operador               | Sítio de ligação do repressor                             |
| (c) Promotor               | Sítio de ligação da RNA polimerase e do complexo CAP-cAMP |
| (d) Gene estrutural $Z$    | Codifica a $\beta$ -galactosidase                         |
| (e) Gene estrutural $Y$    | Codifica a $\beta$ -galactosídeo permease                 |

- 18.5 (a) 1, 2, 3 e 5; (b) 2, 3 e 5.

- 18.7 O mutante  $O^c$  impede a ligação do repressor ao operador. O mutante  $I^s$  não pode se ligar a  $O^c$ . A proteína mutante  $I^s$  tem um defeito no sítio alostérico de ligação da alolactose, mas tem um sítio normal de ligação do operador.

- 18.9 (a)

$$\frac{I^s O^c Z^+ Y^-}{I^s O^c Z^- Y^+}$$

- (b)

$$\frac{I^s O^c Z^+ Y^-}{I^s O^c Z^- Y^+}$$

- 18.11 (a) As mutações  $O^c$  ocupam posição no mapa muito perto do gene estrutural  $Z$ ; as mutações  $I^s$  ocupam posição no mapa um pouco mais distante do gene estrutural (mas ainda muito próximas dele; ver Figura 18.5). (b) Um diploide parcial  $I^s O^c Z^+ Y^- / I^s O^c Z^- Y^+$  apresentaria síntese constitutiva de  $\beta$ -galactosidase e  $\beta$ -galactosídeo permease, enquanto um diploide parcial  $I^s O^c Z^+ Y^- / I^s O^c Z^- Y^+$  seria induzível para síntese dessas enzimas. (c) A mutação  $O^c$  é *cis*-dominante; a mutação  $I^s$  é *trans*-recessiva.
- 18.13 A repressão catabólica evoluiu para garantir o uso de glicose como fonte de carbono, quando esse carboidrato estiver disponível, em lugar de fontes menos eficientes de energia.
- 18.15 Regulação positiva; o complexo CAP-cAMP tem efeito positivo sobre a expressão do operon *lac*. Ele ativa a transcrição dos genes estruturais no operon.
- 18.17 Mecanismos reguladores negativos, como os que incluem o repressor no operon de lactose, bloqueiam a transcrição dos genes estruturais do operon, enquanto mecanismos positivos, como o complexo CAP-cAMP no operon *lac*, promovem a transcrição dos genes estruturais do operon.
- 18.19 A repressão/desrepressão do operon *trp* ocorre no início da transcrição, modulando a frequência com que a RNA polimerase inicia a transcrição a partir dos promotores do operon *trp*. A atenuação modula os níveis de transcrito de *trp* por alteração da frequência de término da transcrição na região líder do operon *trp* (*trpL*).
- 18.21 Primeiro é preciso lembrar que a transcrição e a tradução estão acopladas em procariotos. Quando há triptofano nas células, é produzido tRNA<sup>Trp</sup> carregado com triptofano. Isso possibilita a tradução da sequência líder *trp* ao longo dos dois códons Trp UGG até o códon de término UGA da sequência líder *trp*. Essa tradução da região líder *trp* impede o pareamento de bases entre as sequências líderes de mRNA parcialmente complementares 75-83 e 110-121 (ver Figura 18.15B), o que, por sua vez,



possibilita a formação do “grampo” de término da transcrição com as sequências líderes 110-121 e 126-134 (ver Figura 18.15C).

- 18.23** Tanto a atenuação de *trp* quanto o riborregulador de lisina desativam a expressão gênica por término da transcrição em direção 5' a partir das regiões codificadoras dos genes regulados. Em ambos há formação de estruturas secundárias de mRNA alternativas – com alternância entre a formação de grampos antiterminador e terminador da transcrição – em resposta à presença ou ausência de um metabólito específico (compare a Figura 18.15 e a Figura 2 no texto O futuro: O riborregulador [*riboswitch*] de lisina).

## Capítulo 19

- 19.1** Em eucariotos multicelulares, o ambiente celular é relativamente estável. Não é necessário responder com rapidez a mudanças do ambiente externo. Além disso, o desenvolvimento de um organismo multicelular exige hierarquias reguladoras complexas constituídas de centenas de genes diferentes. A expressão desses genes está submetida à regulação temporal e espacial, com frequência por intrincados processos de sinalização intracelular.
- 19.3** A atividade do gene *dystrophin* poderia ser avaliada por *blotting* do RNA extraído dos diferentes tipos de células e hibridização com uma sonda do gene (*Northern blotting*); outra opção seria a transcrição reversa do RNA em cDNA usando um ou mais iniciadores específicos para o gene *dystrophin*, seguida por amplificação do cDNA resultante por reação em cadeia da polimerase (RT-PCR). Outra técnica seria a hibridização *in situ* – isto é, nas próprias células – do RNA de *dystrophin* com uma sonda do gene. Também seria possível verificar a produção da proteína distrofina por cada tipo celular com uso de anticorpos antidistrofina para analisar proteínas dos diferentes tipos celulares em *Western blots* ou analisar as proteínas nas próprias células – isto é, *in situ*.
- 19.5** Um procedimento seria fornecer às larvas UTP um elemento constitutivo de RNA, marcado radioativamente em diferentes condições – com e sem choque térmico. Depois, preparar amostras de células politênicas dessas larvas para autorradiografia. Se os pufes induzidos por choque térmico contiverem genes transcritos intensamente, o sinal radioativo neles deve ser abundante.
- 19.7** Por recomposição alternativa do transcrito.
- 19.9** Análise por *Northern blot* do RNA extraído de plantas cultivadas com e sem luz, ou amplificação por PCR de cDNA produzido por transcrição reversa desses mesmos extratos de RNA.
- 19.11** Os acentuadores são capazes de agir em qualquer sentido.
- 19.13** Provavelmente não, a menos que o promotor do gene *gfp* seja reconhecido e transcrito pela RNA polimerase de *Drosophila* de modo independente dos elementos de resposta ao choque térmico.
- 19.15** É provável que a mutação seja letal na condição homozigota porque o fator de transcrição controla muitos genes diferentes e é quase certo que uma mutação por mudança de matriz de leitura na sequência codificadora anule a função do fator de transcrição.
- 19.17** O éxon 3 contém um códon de término na matriz de leitura (*in-frame*). Assim, a proteína traduzida do mRNA *Sxl* em machos será mais curta que a proteína traduzida do mRNA *Sxl* mais curto em fêmeas.
- 19.19** O íntron poderia ser introduzido em um vetor de expressão GUS, que seria inserido em plantas *Arabidopsis*. Se o íntron tiver um acentuador que estimula a expressão gênica nas extremidades da raiz, as plantas transgênicas devem ter expressão de GUS nas extremidades da raiz. Veja um exemplo desse tipo de análise no Problema resolvido do Capítulo 19.
- 19.21** Sim. A aparência difusa e inflada indica transcrição intensa dos genes nesse cromossomo – a cromatina está “aberta para negócios”.
- 19.23** Os RNA de interferência curtos têm como alvo moléculas de RNA mensageiro, que não têm íntrons. Assim, o siRNA sintetizado a partir de RNA bifilamentar derivado de um íntron seria ineficaz contra um alvo de mRNA.
- 19.25** O alelo de origem paterna (*b*) será expresso na prole  $F_1$ .
- 19.27** O RNA poderia ser isolado do tecido hepático e encefálico. As técnicas de *Northern blot* ou RT-PCR com esse RNA poderiam determinar qual dos genes (*A* ou *B*) é transcrito em que tecido. No *Northern blot*, as amostras de RNA seriam fracionadas em gel desnaturante e transferidas para uma membrana, depois o RNA na membrana seria hibridizado com sondas gene-específicas, primeiro para um gene, depois para o outro (ou o pesquisador poderia preparar dois *blots* e hibridizar cada um com uma sonda diferente). Na RT-PCR, haveria transcrição reversa das amostras de RNA em cDNA usando iniciadores específicos para cada gene; depois, as moléculas de cDNA seriam amplificadas por PCR padrão, e os produtos das amplificações seriam fracionados por eletroforese em gel para identificar qual RNA gênico estava presente nas amostras originais.
- 19.29** O gene *msl* é inativo em fêmeas.



- 19.31 HP1, a proteína codificada pelo alelo selvagem do gene supressor participa da organização da cromatina. Talvez essa proteína heterocromática se disperse da região perto do ponto de quebra da inversão no cromossomo que tem o alelo *white mottled* e cause a “heterocromatização” do *locus white*. Se houvesse depleção de HP1 por desativação de uma cópia do gene que a codifica – isto é, por mutação supressora do genótipo da mosca, a “heterocromatização” do *locus white* seria menos provável e talvez não ocorresse. Então, o *locus white* seria totalmente ativo em todas as células do olho, produzindo uma cor vermelha uniforme do olho.

## Capítulo 20

- 20.1 A divisão desigual do citoplasma durante as divisões meióticas; o transporte de substâncias de células adjacentes, como as células nutridoras, para o ovócito em *Drosophila*.
- 20.3 Coleção de mutações com fenótipos diagnósticos; mapeamento das mutações e teste de alelismo entre elas; testes de epistasia com mutações em diferentes genes; clonagem de genes individuais e análise de sua função em nível molecular.
- 20.5 Discos imaginais.
- 20.7 Na condição homozigota, a mutação causadora de fenilcetonúria tem efeito materno. As mulheres homozigotas para essa mutação influenciam o desenvolvimento intrauterino dos filhos.
- 20.9 Esterilidade feminina. As fêmeas afetadas por essas mutações põem ovos anormais que não se transformam em embriões viáveis.
- 20.11 As células somáticas que circundam um ovócito de *Drosophila* em desenvolvimento no ovário determinam onde será clivada a proteína spätzle, que é o ligante da proteína receptora Toll. Por fim, essa clivagem ocorrerá na face ventral do embrião em desenvolvimento.
- 20.13 Diferenciação *ey* → *boss* → *sev* → R7
- 20.15 Como o gene *Pax6* produziu o mesmo fenótipo nas moscas que a hiperexpressão do gene *eyeless*, esses genes têm necessariamente homologia funcional e estrutural. Portanto, devem-se esperar olhos adicionais ou primórdios oculares no camundongo com expressão de *eyeless*.
- 20.17 Por teste *Northern blot* de RNA extraído dos tecidos em diferentes períodos durante o desenvolvimento. Hibridiza-se o *blot* com sondas gene-específicas.
- 20.19 A clonagem reprodutiva de mamíferos como carneiros, camundongos e gatos indica que os núcleos de células somáticas têm todas as informações genéticas para guiar o desenvolvimento de um organismo viável completo. Também mostra que mo-

dificações epigenéticas da cromatina, como a inativação do cromossomo X, podem ser redefinidas.

- 20.21 Se cada anticorpo é constituído de um tipo de cadeia leve e um tipo de cadeia pesada, e se as cadeias leves e pesadas podem se combinar livremente, a possibilidade de produzir 100 milhões de anticorpos diferentes implica a existência de 10.000 genes de cadeia leve e 10.000 genes de cadeia pesada ( $10.000 \times 10.000 = 100$  milhões). Se cada cadeia leve tem 220 aminoácidos, cada gene de cadeia leve contém necessariamente  $3 \times 220 = 660$  nucleotídios, porque cada aminoácido é especificado por uma trinca de nucleotídios; do mesmo modo, cada gene de cadeia pesada tem  $3 \times 450 = 1.350$  nucleotídios. Portanto, o genoma tem de conter  $10.000 \times 660 = 6,6$  milhões de nucleotídios dedicados à produção da cadeia leve e  $10.000 \times 1.350 = 13,5$  milhões de nucleotídios dedicados à produção da cadeia pesada. Ao todo, então, o genoma tem de conter 19,5 milhões de nucleotídios dedicados à codificação dos aminoácidos das várias cadeias de anticorpos.

## Capítulo 21

- 21.1 O câncer foi denominado doença genética porque é causado por mutações de genes que controlam o crescimento e a divisão celular. As formas não hereditárias de câncer são causadas por mutações das células somáticas. Essas mutações, porém, podem ser induzidas por fatores ambientais, entre eles fumaça do tabaco, poluentes químicos, radiação ionizante e luz UV. As formas hereditárias de câncer também incluem frequentemente a ocorrência de mutações somáticas induzidas pelo ambiente.
- 21.3 A aneuploidia poderia incluir a perda de cópias funcionais de genes supressores tumorais ou a duplicação imprópria de proto-oncogenes. A perda de genes supressores tumorais eliminaria os freios naturais da divisão celular, e a duplicação de proto-oncogenes aumentaria a abundância de fatores que promovem a divisão celular.
- 21.5 Eles têm íntrons.
- 21.7 Os produtos desses genes têm papéis importantes nas atividades celulares.
- 21.9 As células NIH 3T3 cultivadas provavelmente têm outras mutações que predispõem ao câncer; a transfecção dessas células com um oncogene *c-H-ras* mutante pode ser a última etapa no processo de transformação em células cancerosas. As células embrionárias cultivadas provavelmente não têm as mutações predisponentes necessárias para que se tornem cancerosas; portanto, continuam a se dividir normalmente quando são transfectadas com o oncogene *c-H-ras* mutante.



- 21.11** O retinoblastoma é resultado da condição homozigota para um alelo com perda de função (recessivo). A ocorrência esporádica de retinoblastoma requer duas mutações desse gene na mesma célula ou linhagem celular. Portanto, o retinoblastoma é raro em indivíduos que, no momento da concepção, são homozigotos para o alelo selvagem do gene *RB*. Nesses indivíduos, espera-se que a frequência de tumores nos dois olhos seja igual ao quadrado da frequência de tumores em um olho. Indivíduos heterozigotos para um alelo *RB* mutante necessitam de apenas uma mutação somática para desenvolver retinoblastoma. Como existem milhões de células em cada retina, é alta a probabilidade de que essa mutação somática ocorra em pelo menos uma célula em cada olho, levando ao surgimento de tumores nos dois olhos.
- 21.13** Em nível celular, as mutações com perda de função do gene *RB* são recessivas; uma célula heterozigota para essa mutação divide-se normalmente. Entretanto, quando há uma segunda mutação, essa célula torna-se cancerosa. Se a primeira mutação de *RB* foi hereditária, existe uma alta probabilidade de que o portador dessa mutação desenvolva retinoblastoma porque uma segunda mutação pode ocorrer a qualquer momento durante a formação da retina nos dois olhos. Assim, o indivíduo está predisposto a desenvolver retinoblastoma, e é essa predisposição que tem um padrão de herança dominante.
- 21.15** Por ligação aos fatores de transcrição E2F, pRB impede que esses fatores de transcrição ativem seus genes-alvo, que codificam proteínas que participam do progresso do ciclo celular; portanto, pRB é um regulador negativo de fatores de transcrição que estimulam a divisão celular.
- 21.17** As células homozigotas para uma mutação com perda de função no gene *p16* podem se dividir de maneira descontrolada porque a proteína p16 não seria capaz de inibir a atividade de ciclina-CDK durante o ciclo celular. Portanto, o gene *p16* seria classificado como gene supressor tumoral.
- 21.19** As células homozigotas para uma mutação com perda de função no gene *BAX* seriam incapazes de evitar a repressão da via de morte celular programada pelo produto gênico BCL-2. Consequentemente, essas células seriam incapazes de executar essa via em resposta à lesão do DNA induzida por tratamento com radiação. Essas células continuariam a se dividir e a acumular mutações; por fim, haveria uma boa chance de que se tornassem cancerosas. O gene *BAX* seria classificado como gene supressor tumoral.
- 21.21** Se uma célula fosse heterozigota para uma mutação que causou a união firme e constitutiva de p53 ao DNA de seus genes-alvo, seu crescimento e sua divisão poderiam ser retardados ou poderia haver indução de apoptose. Essa célula seria mais sensível aos efeitos da radiação ionizante porque a radiação aumenta a expressão de p53, e nesse caso p53 estaria predisposta a ativar seus genes-alvo, causando resposta celular vigorosa à radiação.
- 21.23** Provavelmente, elas diminuiriam a capacidade de ligação de pAPC à  $\beta$ -catenina.
- 21.25** A maior irritação do epitélio intestinal causada por uma dieta pobre em fibras e rica em gordura aumentaria a necessidade de divisão celular nesse tecido (para substituir as células perdidas em virtude da irritação), com aumento correspondente da oportunidade de ocorrência de mutações causadoras de câncer.
- 21.27** Não. Aparentemente, há outra via, não mediada por p53, que leva à ativação do gene *p21*.

## Capítulo 22

- 22.1** Alguns genes implicados na cardiopatia são apresentados na Tabela 22.2. Os fatores ambientais incluem alimentação, prática de exercício físico e tabagismo.
- 22.3** A concordância em gêmeos monozigóticos é quase o dobro da concordância em gêmeos dizigóticos. Gêmeos monozigóticos têm duas vezes mais genes em comum que os gêmeos dizigóticos. Os dados são uma forte indicação de que o alcoolismo tem base genética.
- 22.5** Como  $8/2.012$  é aproximadamente  $1/256 = (1/4)^4$ , aparentemente quatro genes determinantes de tamanho eram segregados nos cruzamentos.
- 22.7** Porque  $\Sigma(X_i - \text{média}) = 0$ .
- 22.9**  $3,17/6,08 = 0,52$ .
- 22.11**  $V_g$  é estimada pela média das variâncias das populações endogâmicas:  $9,4 \text{ cm}^2$ .  $V_g$  é estimada pela diferença entre as variâncias da população polinizada aleatoriamente e as populações endogâmicas:  $(26,4 - 9,4) = 17,0 \text{ cm}^2$ . A herdabilidade em sentido amplo é  $H^2 = V_g/V_T = 17,0/26,4 = 0,64$ .
- 22.13** A herdabilidade em sentido amplo tem de ser maior que a herdabilidade em sentido restrito porque  $H^2 = V_g/V_T > V_a/V_T = h^2$ .
- 22.15**  $(15 - 12)(0,3) + 12 = 12,9$  cerdas.
- 22.17**  $h^2 = R/S = (12,5 - 10)/(15 - 10) = 0,5$ ; a seleção para aumento da velocidade de crescimento deve ser eficaz.
- 22.19** Os meios-irmãos têm 25% dos genes em comum. Portanto, o valor máximo de  $h^2$  é de  $0,14/0,25 = 0,56$ .
- 22.21** As correlações para MZJ não são muito diferentes das correlações para MZS. Evidentemente, para essas características de personalidade, a ambientalidade ( $C^2$  na Tabela 22.3) é desprezível.



## Capítulo 23

- 23.1** Frequência de  $L^M$  na população da América Central:  $p = (2 \times 53 + 29) / (2 \times 86) = 0,78$ ;  $q = 0,22$ . Frequência de  $L^M$  na população da América do Norte:  $p = (2 \times 78 + 61) / (2 \times 278) = 0,39$ ;  $q = 0,61$ .
- 23.3**  $q^2 = 0,0004$ ;  $q = 0,02$ .
- 23.5** Frequência de pessoas que sentem o sabor (genótipos  $TTe$  e  $Tt$ ):  $(0,4)^2 + 2(0,4)(0,6) = 0,64$ . Frequência de pessoas  $TT$  que sentem o sabor entre todos os que são sensíveis ao sabor:  $(0,4)^2 / (0,64) = 0,25$ .
- 23.7**  $(0,00025)^2 = 6,25 \times 10^{-8}$ .
- 23.9** Em mulheres, a frequência do fenótipo dominante é 0,36. A frequência do fenótipo recessivo é 0,64 =  $q^2$ ; assim,  $q = 0,8$  e  $p = 0,2$ . Portanto, a frequência do fenótipo dominante em homens é  $p = 0,2$ .
- 23.11** A frequência de heterozigotos =  $H = 2pq = 2p(1 - p)$ . Usando o cálculo, determine a derivada de  $H$  e iguale o resultado a zero para determinar o valor de  $p$  que maximiza  $H$ :  $dH/dp = 2 - 4p = 0$  significa que  $p = 2/4 = 0,5$ .
- 23.13** Supondo que a população esteja em equilíbrio de Hardy-Weinberg, a frequência do alelo para cor clara é a raiz quadrada da frequência de homozigotos recessivos. Assim,  $q = \sqrt{0,49} = 0,7$ , e a frequência do alelo para cor escura é  $1 - q = p = 0,3$ . De  $p^2 = 0,09$ , estimamos que  $0,09 \times 100 = 9$  das mariposas escuras na amostra são homozigotas para o alelo dominante.
- 23.15** A frequência final de  $GG$  é 0,2; a frequência final de  $gg$  é 0,8.
- 23.17** (a) A frequência de  $A$  na população unificada é 0,5 e a frequência de  $a$  também é 0,5; (b) 0,25 ( $AA$ ), 0,50 ( $Aa$ ) e 0,25 ( $aa$ ); (c) persistirão as frequências em (b).
- 23.19** As aptidões relativas podem ser calculadas por divisão de cada probabilidade de sobrevivência pela maior probabilidade (0,92). Assim, as aptidões relativas são 1 para  $GG$ ,  $0,98 = 1 - 0,02$  para  $Gg$ , e  $0,61 = 1 - 0,39$  para  $gg$ . Os coeficientes de seleção são  $s_1 = 0,02$  para  $Gg$  e  $s_2 = 0,39$  para  $gg$ .
- 23.21** (a) Use o seguinte esquema:

| Genótipo                                     | $CC$                     | $Cc$                     | $cc$                |
|----------------------------------------------|--------------------------|--------------------------|---------------------|
| Frequência de Hardy-Weinberg                 | $(0,98)^2 = 0,9604$      | $2(0,98)(0,02) = 0,0392$ | $(0,02)^2 = 0,0004$ |
| Aptidão relativa                             | 1                        | 1                        | 0                   |
| Contribuição relativa para a próxima geração | $(0,9604) \times 1$      | $(0,0392) \times 1$      | 0                   |
| Contribuição proporcional                    | $0,9604/0,9996 = 0,9608$ | $0,0392/0,9996 = 0,0392$ | 0                   |

A nova frequência do alelo para fibrose cística é  $(0,5)(0,0392) = 0,0196$ ; assim, a incidência da doença será

$(0,0196)^2 = 0,00038$ , que é apenas ligeiramente menor que a incidência na geração anterior. (b) A incidência de fibrose cística não se modifica muito porque a seleção só pode atuar contra o alelo recessivo quando está em homozigotos, que são raros na população.

- 23.23**  $q^2 = 4 \times 10^{-5}$ ; assim  $q = 6,3 \times 10^{-3}$  e  $2pq = 0,0126$ .
- 23.25** A probabilidade de fixação de  $A_2$  é 0,5; a probabilidade de perda de  $A_3$  é  $1 - 0,3 = 0,7$ .
- 23.27**  $p = 0,2$ ; em equilíbrio,  $p = t / (s + t)$ . Como  $s = 1$ , podemos calcular  $t$ ;  $t = 0,25$ .
- 23.29** No equilíbrio mutação-seleção,  $q = \sqrt{u/s} = q = \sqrt{10^{-6}/1} = 0,001$ .

## Capítulo 24

- 24.1** Entre outras coisas, Darwin observou espécies nas ilhas diferentes entre si e das espécies continentais, mas que ainda eram suficientemente semelhantes para indicar um parentesco. Ele também observou variação nas espécies, sobretudo em raças domesticadas, e viu como as características de um organismo poderiam ser modificadas por cruzamento seletivo. As observações de organismos fossilizados indicaram que algumas espécies foram extintas.
- 24.3** A frequência do alelo  $a$  é 0,06 na população da África do Sul e 0,42 na população da Inglaterra. As frequências genotípicas previstas considerando-se que haja cruzamento aleatório são:

| Genótipo | África do Sul          | Inglaterra             |
|----------|------------------------|------------------------|
| $aa$     | $(0,06)^2 = 0,004$     | $(0,42)^2 = 0,18$      |
| $ab$     | $2(0,06)(0,94) = 0,11$ | $2(0,42)(0,58) = 0,49$ |
| $bb$     | $(0,94)^2 = 0,88$      | $(0,58)^2 = 0,33$      |

- 24.5** Na amostra, a frequência do  $F$  é  $(2 \times 32 + 46) / (2 \times 100) = 0,55$  e a frequência do alelo  $S$  é  $1 - 0,55 = 0,45$ . As frequências genotípicas previstas e observadas são:

| Genótipo | Observado | Previsão de Hardy-Weinberg        |
|----------|-----------|-----------------------------------|
| $FF$     | 32        | $100 \times (0,55)^2 = 30,25$     |
| $FS$     | 46        | $100 \times 2(0,55)(0,45) = 49,5$ |
| $SS$     | 22        | $100 \times (0,45)^2 = 20,25$     |

Para verificar o consenso entre os valores observado e previsto, calculamos uma estatística de qui-quadrado com um grau de liberdade:  $\chi^2 = \sum (\text{obs.} - \text{prev.})^2 / \text{prev.} = 0,50$ , que não é significativa no nível de 5%. Assim, a população parece estar em equilíbrio de Hardy-Weinberg para o locus da álcool desidrogenase.

- 24.7** Na terceira posição de alguns dos códons. Por causa da degeneração do código genético, diferentes códons podem especificar o mesmo aminoácido.



A degeneração é mais acentuada na terceira posição de muitos códons, onde pode haver diferentes nucleotídeos sem modificação do aminoácido especificado.

- 24.9** Os carboidratos complexos não são “documentos da história evolutiva” porque, embora sejam polímeros, são tipicamente constituídos de uma subunidade incorporada repetidas vezes em uma cadeia. Esse polímero tem pequeno ou nenhum “conteúdo de informações”. Assim, a oportunidade de distinguir um carboidrato complexo obtido de dois organismos diferentes é pequena ou nula. Além disso, os carboidratos complexos não são parte do mecanismo genético; sua formação é, em última análise, especificada pela ação de enzimas, que são produtos gênicos, mas eles próprios não são material genético nem produtos do material genético.
- 24.11** As histidinas são rigorosamente conservadas porque têm uma importante função – ancoragem do grupo heme na hemoglobina. Como esses aminoácidos são fortemente restringidos pela seleção natural, não evoluem por mutação e deriva genética aleatória.
- 24.13** Estime o número médio de substituições por sítio na molécula de ribonuclease por  $-\ln(S)$ , onde  $S = (124 - 40)/124 = 0,68$ , a proporção de aminoácidos iguais nas moléculas de rato e de boi. Portanto, o número médio de substituições por sítio desde que as linhagens de boi e rato divergiram de um ancestral comum é 0,39. A taxa evolutiva nas linhagens de boi e rato é de  $0,39/(2 \times 80 \text{ milhões de anos}) = 2,4 \text{ substituições por sítio a cada bilhão de anos}$ .
- 24.15** O inverso da taxa, isto é,  $1/K$ .
- 24.17** A proteína com a maior taxa evolutiva não é tão restringida pela seleção natural como a proteína com a menor taxa evolutiva.
- 24.19** As sequências repetitivas próximas podem mediar o pareamento deslocado durante a meiose. A troca com a participação das sequências deslocadas pode duplicar a região entre elas.
- 24.21** Cruzamento de *D. mauritiana* com *D. simulans* para verificar se essas duas espécies estão isoladas reprodutivamente. Por exemplo, elas produzem descendentes? Em caso afirmativo, a prole é fértil?
- 24.23** A interação *Kpn-pn* é um exemplo do tipo de epistasia negativa que poderia impedir que populações evoluídas separadamente se fundissem em uma população panmítica. A mutação *Kpn* teria evoluído em uma população e uma mutação *pn* em outra população geograficamente separada. Quando as populações se juntam, as duas mutações podem ser reunidas na mesma mosca por intercruzamento. Se a combinação dessas mutações for letal, as populações previamente separadas não serão capazes de trocar genes; isto é, elas estarão reprodutivamente isoladas.



# Glossário

Este glossário introduz alguns termos básicos e recorrentes no livro. Nomes de substâncias químicas, definições de termos especializados e variantes de nomes básicos foram omitidos, mas constam do índice remissivo. Procure no índice remissivo os termos que não encontrar no glossário.

## A

- Aberração cromossômica.** Estrutura ou número de cromossomos anormal; inclui deficiência ou deleção, duplicação, inversão, translocação, aneuploidia, poliploidia ou qualquer outra variação do padrão normal.
- Abscissa.** A escala horizontal de um gráfico.
- Acentuador.** Substância ou objeto que intensifica uma atividade química ou um processo fisiológico; um gene de efeito maior ou modificador que intensifica um processo fisiológico; uma sequência de DNA que influencia a transcrição de um gene adjacente.
- Ácido desoxirribonucleico.** Veja DNA.
- Ácido nitroso.**  $\text{HNO}_2$ , um potente mutágeno químico.
- Ácido nucleico.** Macromolécula constituída de ácido fosfórico, pentose e bases orgânicas; DNA e RNA.
- Ácido ribonucleico.** Veja RNA.
- Acoplamento (configuração *cis*).** Condição na qual um heterozigoto duplo recebeu duas mutações ligadas de um genitor e seus alelos selvagens do outro (p. ex.,  $a b/a b \times +/+ +$  produz  $a b/+ +$  (cf. **Repulsão**).
- Acridinas.** Classe de moléculas policíclicas com carga elétrica positiva que são intercaladas no DNA e induzem mutações por mudança da matriz de leitura; usadas como corantes.
- Acrocêntrico.** Termo modificador que designa um cromossomo ou cromátide cujo centrômero está perto da extremidade.
- Activator (Ac).** Elemento transponível (transpósion) do milho que codifica uma enzima transposase capaz de catalisar o movimento de elementos Ac e de outros membros da família Ac/Ds.
- ADA-SCID (imunodeficiência combinada grave com deficiência de adenosina desaminase).** Distúrbio autossômico recessivo em seres humanos causado pela ausência da enzima adenosina desaminase, que catalisa a degradação da desoxiadenosina. Na ausência dessa enzima, os derivados tóxicos desse nucleosídeo acumulam-se e destroem as células necessárias para a resposta imune normal a infecções.
- Adaptação.** Ajuste de um organismo ou de uma população a um ambiente.
- Adenina (A).** Uma base purínica encontrada no RNA e no DNA.
- Adição gênica.** Acréscimo de cópia funcional de um gene ao genoma de um organismo.
- A-DNA.** Dupla hélice de DNA dextrógira que tem 11 pares de bases por volta. O DNA apresenta-se nessa forma quando parcialmente desidratado.
- Agente hidroxilante.** Substância química – como o mutágeno hidroxilamina – que transfere grupos hidroxila para outras moléculas.
- Agente intercalar.** Substância química capaz de se inserir entre pares de base adjacentes em uma molécula de DNA.
- Agentes alquilantes.** Substâncias químicas que transferem grupos alquila (metila, etila e assim por diante) para as bases no DNA.
- AIDS (síndrome de imunodeficiência adquirida).** Doença humana geralmente fatal na qual o sistema imune é destruído pelo vírus da imunodeficiência humana (HIV).
- Albinismo.** Ausência de pigmento na pele, nos pelos e nos olhos de um animal. Ausência de clorofila em vegetais.
- Alça t.** Alça de DNA formada por sequências repetidas de telômeros na extremidade de um cromossomo linear quando um filamento único na terminação 3' invade uma unidade de repetição na direção 5' e emparelha-se com o filamento complementar ao mesmo tempo que desloca o filamento equivalente.
- Alcaptonúria.** Distúrbio metabólico hereditário. As pessoas com alcaptonúria excretam quantidade excessiva de ácido homogentísico (alcaptona) na urina.
- Alças R.** Regiões de DNA unifilamentares em híbridos de RNA-DNA formados *in vitro* em condições em que os filamentos duplos de RNA-DNA são mais estáveis que os filamentos duplos de DNA-DNA.
- Alelo (alelomorfo; *adj.*, alélico, alelomórfico).** Elemento de um par, ou série, de formas alternativas de um gene que ocorrem em determinado *locus* de um cromossomo. Os alelos são indicados pelo mesmo símbolo básico (p. ex., *D* para pés de ervilha altos e *d* para pés de ervilha anões). (Veja também **Alelos múltiplos**.)
- Alelo nulo.** Forma mutante de um gene que não cria um produto ou que cria um produto totalmente inativo.
- Alelos codominantes.** Alelos que produzem efeitos independentes quando heterozigotos.
- Alelos múltiplos.** Condição em que determinado gene ocorre em três ou mais formas alélicas em uma população de organismos.
- Aleurona.** Camada externa do endosperma de uma semente.
- Alongamento (da síntese de DNA, RNA ou proteínas).** Incorporação da segunda subunidade e das subsequentes (nucleotídios e aminoácidos) durante a síntese de uma macromolécula (DNA, RNA ou polipeptídio).
- Alopoliploide.** Poliploide que tem conjuntos de cromossomos de diferentes espécies; poliploide que contém conjuntos



- de cromossomos geneticamente diferentes oriundos de duas ou mais espécies.
- Alotetraploide.** Organismo com quatro genomas derivado da hibridização de diferentes espécies. Em geral, as formas que se estabelecem têm dois genomas de uma espécie e os dois outros de outra espécie.
- Alozima.** Variante de uma enzima detectada por eletroforese.
- Ambientalidade.** A proporção da variância fenotípica total de uma característica quantitativa decorrente dos efeitos de um ambiente em comum.
- Ambiente.** Reunião de todas as condições externas e influências que afetam a vida e o desenvolvimento de um organismo.
- Aminoácido.** Qualquer substância orgânica pertencente a uma classe que contém um grupo amino ( $\text{NH}_2$ ) e um grupo carboxila ( $\text{COOH}$ ). Os aminoácidos são os precursores das proteínas. Alanina, prolina, treonina, histidina, lisina, glutamina, fenilalanina, triptofano, valina, arginina, tirosina e leucina estão entre os aminoácidos mais comuns.
- Aminoacil-tRNA sintetases.** Enzimas que catalisam a formação de ligações de alta energia entre aminoácidos e moléculas de tRNA.
- Âmnio.** Membrana delgada que reveste a bolsa cheia de líquido na qual se desenvolve o embrião de vertebrados superiores.
- Amniocentese.** Procedimento de coleta de líquido amniótico de uma mulher grávida. A análise direta da composição química do líquido é realizada para diagnóstico de algumas doenças. As células são cultivadas, e examinam-se os cromossomos em metáfase para identificação de irregularidades (p. ex., trissomia).
- Amórfico.** Termo aplicado a um alelo mutante que suprime totalmente a expressão gênica. Esse alelo mutante é denominado amorfo.
- Amostra.** Grupo de itens selecionados para representar uma grande população.
- AMP cíclico.** 3', 5'-monofosfato de adenosina, uma pequena molécula que precisa se ligar à proteína ativadora do catabolismo (CAP) para que o complexo (CAP/cAMP) se ligue aos promotores de óperons e estimule a transcrição.
- Amplificação (moléculas de DNA recombinantes).** Produção de muitas cópias de uma molécula de DNA recombinante recém-construída.
- Amplificação do éxon.** Procedimento usado para identificar regiões codificadoras (éxons) ladeadas por sítios de *splicing* 5' e 3' de íntron.
- Amplificação gênica.** Fenômeno pelo qual o DNA de um gene ou de um conjunto de genes específico é replicado separado do restante do genoma para aumentar o número de cópias do gene.
- Anáfase.** Estágio da mitose ou meiose em que os cromossomos-filhos vão da placa equatorial até polos opostos da célula (em direção às extremidades do fuso). A anáfase sucede a metáfase e precede a telófase.
- Anáfase I.** Estágio da primeira divisão meiótica em que os cromossomos homólogos duplicados separam-se e começam a se deslocar em direção a polos opostos da célula.
- Anáfase II.** Estágio da segunda divisão meiótica em que as cromátides-irmãs de um cromossomo duplicado se separam e começam a se deslocar em direção a polos opostos da célula.
- Análise do perfil de DNA (análise da impressão digital de DNA).** Uso de dados da sequência de DNA – sobretudo de repetições curtas em série (STR) altamente polimórficas e repetições em série em número variável (VNTR) – para identificação pessoal.
- Análogos de bases.** Bases purínicas ou pirimidínicas não naturais, um pouco diferentes das bases normais, que podem ser incorporadas aos ácidos nucleicos. Muitas vezes são mutagênicas.
- Androgênio.** Hormônio masculino que controla a atividade sexual em animais vertebrados.
- Anemia.** Distúrbio caracterizado por palidez, fraqueza e dispnéia, decorrente da deficiência de hemoglobina ou diminuição do número de hemácias.
- Aneuploide.** Organismo ou célula cujo número de cromossomos não é um múltiplo exato do monoploide ( $n$ ) com um genoma, isto é, hiperploide, maior (p. ex.,  $2n + 1$ ), ou hipoploide, menor (p. ex.,  $2n - 1$ ). Também designa os casos de duplicação ou ausência de parte de um cromossomo.
- Anfidiploide.** Espécie ou tipo de vegetal derivado da duplicação dos cromossomos na  $F_1$  híbrida de duas espécies; um aloploidio. No anfidiploide, as duas espécies são conhecidas, enquanto nos outros aloploidios elas podem ser desconhecidas.
- Antera.** Órgão das flores que produz pólen.
- Anticódon.** Três bases de um RNA transportador que são complementares às três bases de um códon específico no RNA mensageiro.
- Anticorpo.** Substância de um tecido ou líquido do corpo que combate uma substância estranha (antígeno).
- Antígeno.** Substância, geralmente uma proteína, à qual se liga um anticorpo ou um receptor de células T quando introduzida em um organismo vertebrado.
- Antiterminador de transcrição.** Proteína que impede a RNA polimerase de parar a transcrição em sequências específicas de término de transcrição.
- Apomixia.** Método assexuado de reprodução em que há produção de oosferas não reduzidas (geralmente diploides) que se desenvolvem sem fertilização.
- Apoptose.** Fenômeno no qual células eucariotas morrem por eventos internos geneticamente programados.
- Aptidão relativa.** A sobrevivência e a capacidade reprodutiva de um genótipo em uma população em comparação com a sobrevivência e a capacidade reprodutiva de outro genótipo nessa população.
- Aptidão.** O número de filhos de um indivíduo, muitas vezes comparado com a média da população ou com algum outro padrão, como o número de filhos de determinado genótipo.
- Asco.** Saco reprodutivo do estágio sexuado de um tipo de fungo (ascomiceto) no qual são produzidos os ascósporos.
- Ascósporo.** Um dos esporos contidos no asco de determinados fungos como *Neurospora*.
- Assinapse.** Ausência de pareamento ou pareamento parcial de cromossomos homólogos durante a prófase meiótica.
- Atenuação.** Mecanismo de controle da expressão gênica em procariotos que implica o término prematuro da transcrição.
- Atenuador.** Sequência de nucleotídeos na região 5' de um gene procarioto (ou em seu RNA) que causa o término prematuro da transcrição, possivelmente por formação de uma estrutura secundária.



**Ativador (da expressão gênica).** Produto de genes reguladores que estimula, ou ativa, a expressão de outros genes.

**ATP.** Trifosfato de adenosina: substância rica em energia que promove determinadas atividades celulares.

**Autofertilização.** Processo pelo qual o pólen de determinado vegetal fertiliza os óvulos do mesmo vegetal. Diz-se que os vegetais fertilizados dessa forma foram autopolinizados. Processo análogo ocorre em alguns animais, como nematódeos e moluscos.

**Autônomo.** Termo usado para designar toda unidade biológica que tem função independente, isto é, sem a ajuda de outra unidade. Por exemplo, um elemento transponível que codifica uma enzima para sua própria transposição (cf. **Não autônomo**).

**Autopoliploide.** Poliploide que tem conjuntos múltiplos e idênticos ou quase idênticos de cromossomos (genomas). Uma espécie poliploide com genomas derivados da mesma espécie original.

**Autorradiografia.** Registro fotográfico no qual se marca uma substância, como o DNA, com uma substância radioativa, como a timidina tritiada, e se espera determinado período até que a imagem produzida por decaimento radioativo seja revelada em um filme.

**Autossomo.** Qualquer cromossomo não sexual.

**Auxotrófico.** Microrganismo mutante (p. ex., bactéria ou levedura) que não cresce em meio mínimo, mas requer o acréscimo de alguma substância como um aminoácido ou uma vitamina.

## B

**BAC (cromossomos artificiais bacterianos).** Vetores de clonagem construídos a partir de fatores de fertilidade bacteriana (F); como os vetores YAC, eles aceitam insertos grandes de 200 a 500 kb.

**Bactérias lisogênicas.** Bactérias que abrigam bacteriófagos temperados.

**Bacteriófago.** Vírus que ataca bactérias. Esses vírus são denominados bacteriófagos porque destroem as bactérias hospedeiras.

**Banda satélite.** Banda formada por DNA em gradiente de densidade diferente e menor que o DNA da banda principal. A banda satélite contém sequências de DNA repetidas denominadas DNA satélites com densidade menor ou maior que o DNA da banda principal.

**Bandeamento cromossômico.** Coloração de cromossomos que mostra áreas claras e escuras ao longo dos cromossomos. As comparações laterais identificam os pares. É possível identificar cada cromossomo humano pelo padrão de bandas.

**B-DNA.** DNA bifilamentar cuja hélice é dextrógira, com 10,4 pares de bases por volta; a conformação do DNA quando presente em soluções aquosas com baixa concentração de sal.

**Biblioteca de cDNA.** Conjunto de clones de cDNA contendo cópias dos RNA isolados de um organismo ou de um tipo específico de tecido ou célula de um organismo.

**Biblioteca de DNA genômico.** Coleção de genes contendo as sequências do DNA genômico de um organismo.

**Bioinformática.** Estudo da genética e de outras informações biológicas usando técnicas de computador e estatísticas.

**Biometria.** Aplicação de métodos estatísticos ao estudo de problemas biológicos.

**Biopsia de vilosidades coriônicas.** Procedimento no qual amostras de células de um embrião são coletadas para teste genético.

**Bivalente.** Par de cromossomos homólogos em sinapse ou associados que passaram por um processo de duplicação para formar um grupo de quatro cromátides.

**Blastômero.** Qualquer uma das células formadas nas primeiras clivagens no desenvolvimento de animais.

**Blástula.** Em animais, forma embrionária inicial que sucede o estágio de mórula; em geral, uma camada simples ou uma esfera de células.

**Bolha de replicação.** Região localizada de separação do filamento complementar que ocorre na origem da replicação no início da replicação do DNA.

**Bolha de transcrição.** Segmento de DNA com desenrolamento local no qual está sendo sintetizado um transcrito de RNA.

**Bombardeamento com microprojéteis.** Procedimento que visa transformar células vegetais pelo disparo de partículas de tungstênio ou de ouro revestidas de DNA para dentro das células.

## C

**Cadeia kappa.** Uma das duas classes de cadeias leves de anticorpo (cf. **Cadeia lambda**).

**Cadeia lambda.** Uma das duas classes de cadeias leves de anticorpo (cf. **Cadeia kappa**).

**Calvície de padrão masculino.** Forma hereditária de calvície na qual a rarefação do cabelo começa no topo da cabeça.

**Caminhada no cromossomo (*chromosome walking*).** Procedimento que usa grandes clones superpostos para o deslocamento sequencial de um sítio para outro ao longo de um cromossomo. (Veja também **Clonagem posicional**.)

**Câncer colorretal não polipóide.** Tipo de câncer encontrado na porção inferior do trato digestivo, às vezes herdado como um distúrbio dominante.

**Cap 5' (mRNA).** O *cap* (capacete) de 7-metil-guanosina acrescentado à maioria dos mRNA eucarióticos após a transcrição.

**Cápsulas de polissacarídio.** Revestimentos de carboidrato com especificidade antigênica que estão presentes em alguns tipos de bactérias.

**Característica de limiar.** Característica que tem manifestação descontínua, mas é influenciada por variações genéticas e ambientais contínuas.

**Característica multifatorial.** Característica determinada por uma combinação de vários fatores genéticos e ambientais.

**Características quantitativas.** Fenótipos mensuráveis, como altura, peso e velocidade de crescimento.

**Caráter (característica).** Um dos muitos detalhes da estrutura, forma, substância ou função que compõem um organismo individual.

**Carboidrato.** Molécula constituída de carbono, hidrogênio e oxigênio na proporção de 1:2:1; uma molécula de açúcar ou uma macromolécula composta de subunidades de açúcar.

**Carcinógeno.** Agente capaz de induzir câncer em um organismo.



- Cariótipo.** A constituição cromossômica de uma célula ou um indivíduo; os cromossomos dispostos em ordem de comprimento e de acordo com a posição do centrômero; também é a fórmula abreviada da constituição cromossômica, como 47, XX + 21 na trissomia do 21 em ser humano.
- Cauda poli(A) (mRNA).** Trecho de poliadenosina com 20 a 200 nucleotídeos de comprimento que é acrescentado às extremidades 3' da maioria dos mRNA eucarióticos após a transcrição.
- cDNA (DNA complementar).** Molécula de DNA sintetizada *in vitro* a partir de um molde de RNA.
- Célula (ou núcleo) totipotente.** Célula (ou núcleo) indiferenciada como um blastômero que, quando isolado ou adequadamente transplantado, pode dar origem a um embrião completo.
- Célula doadora.** Bactéria que doa DNA para outra célula (receptora) durante a recombinação em bactérias (cf. **Célula receptora**).
- Célula F<sup>+</sup>.** Bactéria que contém um fator de fertilidade (F) autônomo. Veja **Fator F**.
- Célula germinativa.** Célula reprodutiva capaz, quando madura, de ser fertilizada e reproduzir um organismo inteiro (cf. **Célula somática**).
- Célula receptora.** Bactéria que recebe DNA de outra célula (doadora) durante a recombinação (cf. **Célula doadora**).
- Célula somática.** Célula constituinte do corpo, ao contrário de uma célula germinativa que é capaz, quando fertilizada, de reproduzir o organismo.
- Célula-filha.** Produto da divisão celular.
- Célula-mãe.** Célula preparada para se dividir por mitose ou meiose.
- Células do polo.** Grupo de células na porção posterior de embriões de *Drosophila* que são precursoras da linhagem germinativa do adulto.
- Células eucarióticas.** As células dos organismos classificados como eucariotos. Essas células são caracterizadas por terem um núcleo, que contém o DNA cromossômico, delimitado por membrana.
- Células procarióticas.** As células dos organismos classificados como procariotos. Essas células são caracterizadas por não terem um núcleo, delimitado por membrana, que contém o DNA cromossômico.
- Células T auxiliares.** Células T cuja resposta a um antígeno exibido por um macrófago é a estimulação da transformação de linfócitos B e T em plasmócitos produtores de anticorpos e células T citotóxicas, respectivamente.
- Células-tronco embrionárias (CTE).** Células presentes em embriões capazes de se diferenciar em muitos tipos diferentes de tecidos e/ou órgãos.
- Célula-tronco.** Célula capaz de extensa proliferação e cuja prole é capaz de se diferenciar em tipos celulares especializados.
- CentiMorgan.** Veja **Unidade de crossing over**.
- Centrifugação de equilíbrio por gradiente de densidade.** Procedimento usado para separar macromoléculas com base em sua densidade (massa por unidade de volume).
- Centríolo.** Organela existente em muitas células animais e que aparentemente participa da formação do fuso durante a mitose.
- Centro organizador de microtúbulos (MTOC).** Região de uma célula eucariótica que produz os microtúbulos usados durante a divisão celular. Em células de animais, o MTOC está associado a organelas distintas denominadas centrosomos.
- Centrômero.** Região de um cromossomo onde há fixação à fibra do fuso mitótico.
- Centrossomo.** Organela em forma de barril associada ao fuso mitótico em células animais.
- Chaperona.** Proteína que auxilia novos polipeptídeos a se dobrarem em suas estruturas tridimensionais próprias.
- Chip de DNA.** Veja **Chip gênico**.
- Chip gênico.** Uma pequena bolacha (*wafer*) de silício ou outro suporte sólido que contém uma grande quantidade de oligonucleotídeos ou sondas de hibridização de cDNA dispostos em sua superfície em um padrão específico, ou microarranjo.
- Ciclo celular.** Eventos cíclicos ocorridos durante as divisões das células mitóticas. O ciclo celular oscila entre mitose e interfase, que é dividida em G<sub>1</sub>, S e G<sub>2</sub>.
- Cílio.** Estrutura locomotora, semelhante a um pelo, presente em determinadas células; estrutura locomotora de protozoários ciliados.
- Cinética.** Processo dinâmico que inclui movimento.
- Cinetócoro.** Estrutura proteinácea associada ao centrômero de um cromossomo durante a divisão celular eucariótica; o ponto ao qual se fixam os microtúbulos para deslocar o cromossomo durante o processo de divisão.
- Citocinese.** Divisão citoplasmática e outras alterações exclusivas da divisão nuclear que são parte da mitose ou meiose.
- Citoesqueleto.** Sistema complexo de fibras e filamentos que proporciona sustentação às células e participa do deslocamento de componentes celulares pelo citoplasma.
- Citogenética.** Área da biologia relacionada com os cromossomos e suas implicações na genética.
- Citologia.** Estudo da estrutura e da função das células.
- Citoplasma.** Protoplasma de uma célula fora do núcleo, no qual residem as organelas celulares (mitocôndrias, plastídios e semelhantes); todas as partes vivas da célula, exceto o núcleo.
- Citosina (C).** Base pirimidínica encontrada no RNA e no DNA.
- Classe modal.** Em uma distribuição de frequência, a classe que tem a maior frequência.
- Clonagem (gene).** Produção de muitas cópias de um gene ou de uma sequência específica de DNA.
- Clonagem gênica.** Incorporação de um gene de interesse em uma molécula de DNA autorreplicante e amplificação da molécula de DNA recombinante resultante em uma célula hospedeira apropriada.
- Clonagem posicional.** Isolamento de um clone de um gene ou outra sequência de DNA com base na posição no mapa no genoma.
- Clonagem reprodutiva.** Processo no qual o núcleo de um gameta feminino é substituído pelo núcleo de uma célula de um organismo desenvolvido com o objetivo de produzir um novo organismo geneticamente idêntico ao doador.
- Clonagem terapêutica.** Processo no qual o núcleo do ovócito é substituído pelo núcleo de uma célula de doador (possivelmente diferenciada) com o objetivo de produzir uma po-



pulação de células-tronco com genótipo igual ao da célula doadora. Essas células-tronco poderiam, então, ser usadas para substituir células perdidas no organismo doador.

**Clone.** Todos os indivíduos originados de propagação vegetativa de um único indivíduo original. Em biologia molecular, uma população de moléculas de DNA idênticas que apresentam uma sequência de DNA específica de um organismo.

**Cloroplasto.** Organela verde no citoplasma de vegetais que contém clorofila e na qual é sintetizado o amido. Um mecanismo de herança citoplasmática, independente de genes nucleares, foi associado a essas organelas citoplasmáticas.

**Código genético.** O conjunto de 64 trinucleotídeos que especificam os 20 aminoácidos e o início e o término da cadeia polipeptídica.

**Códon de iniciação.** Sequência de três nucleotídeos no mRNA – geralmente AUG, às vezes GUG – que indica o início de um novo polipeptídio durante a tradução.

**Códon de término de cadeia.** Códon que especifica o término da cadeia polipeptídica em vez da incorporação de um aminoácido. Há três códons desse tipo (UAA, UAG e UGA), reconhecidos por fatores de liberação de proteínas, e não por tRNA.

**Códon.** Conjunto de três nucleotídeos adjacentes em uma molécula de mRNA que especifica a incorporação de um aminoácido a uma cadeia polipeptídica ou que indica o término da síntese do polipeptídio. Códon com essa última função são denominados códons de término ou finalizadores.

**Coefficiente binomial.** Termo que indica o número de maneiras de obter os dois resultados possíveis em uma experiência na qual só são possíveis dois resultados.

**Coefficiente de coincidência.** A razão entre a frequência observada e a frequência esperada de *crossing overs* duplos, calculada com base na suposição de que os *crossing overs* em segmentos adjacentes do cromossomo são independentes.

**Coefficiente de endogamia.** Probabilidade de que dois alelos em um indivíduo sejam idênticos por descendência de um ancestral comum.

**Coefficiente de relação.** Fração de genes comuns a dois indivíduos em virtude de ascendência comum.

**Coefficiente de seleção.** Número que mede a aptidão de um genótipo em relação a um padrão.

**Coefficiente.** Número que expressa a quantidade de alguma variação ou efeito em determinadas condições (p. ex., o coeficiente de endogamia).

**Coenzima.** Substância necessária para a atividade de uma enzima.

**Coincidência.** Razão entre a frequência observada e a frequência esperada de *crossing overs* duplos, na qual a frequência esperada é calculada com base na suposição de que os dois eventos de *crossing over* são independentes um do outro.

**Cointegrado.** Molécula de DNA formada pela fusão de duas moléculas de DNA diferentes, geralmente mediada por um elemento transponível.

**Colchicina.** Alcaloide derivado da *Colchicum autumnale* (planta dama-nua) e usado para interromper a formação do fuso e a mitose.

**Colinearidade (adj., colinear).** Relação na qual as unidades de uma molécula ocorrem na mesma sequência que as uni-

dades de outra molécula que especificam; por exemplo, os nucleotídeos de um gene são colineares em relação aos aminoácidos do polipeptídio codificado por ele.

**Colônia.** Coleção compacta de células produzida pela divisão de uma única célula progenitora.

**Compensação de dose.** Fenômeno no qual a atividade de um gene é aumentada ou diminuída de acordo com o número de cópias desse gene na célula.

**Competência (adj. competente).** Capacidade que tem uma célula bacteriana de incorporar DNA e sofrer transformação genética.

**Complementaridade.** Relação entre os dois filamentos de uma dupla hélice de DNA. A timina em um filamento faz par com a adenina no outro filamento, e a citosina em um filamento faz par com a guanina no outro filamento.

**Complexo de Golgi.** Sistema de membranas no interior das células que participa da secreção de substâncias celulares.

**Complexo de silenciamento induzido por RNA (RISC).** Complexo proteico que usa RNA bifilamentar para produzir e direcionar pequenos RNA de interferência para RNA mensageiros complementares nas células eucarióticas.

**Complexo sinaptonêmico.** Estrutura semelhante a uma fita, formada entre homólogos em sinapse no fim da primeira prófase meiótica, que une as cromátides ao longo de sua extensão e facilita a troca de cromátides.

**Configuração cis.** Veja **Acoplamento**.

**Configuração trans.** Veja **Repulsão**.

**Conídio.** Esporo assexuado produzido por hifa especializada em determinados fungos.

**Conjugação.** União de células sexuais (gametas) ou de organismos unicelulares durante a fertilização; em *Escherichia coli*, transferência unidirecional de material genético de um doador (célula “masculina”) para um receptor (célula “feminina”).

**Consanguinidade.** Relação decorrente de descendência de um ancestral comum.

**Contíg.** Conjunto de clones superpostos que constituem um mapa físico de uma porção de um cromossomo.

**Conversão gênica.** Processo, muitas vezes associado à recombinação, durante o qual há replicação de um alelo à custa de outro, o que ocasiona razões de segregação não mendelianas. Nas tétrades completas, por exemplo, a razão pode ser de 6:2 ou 5:3 em vez da razão esperada de 4:4.

**Copolímeros.** Misturas constituídas de mais de um monômero; por exemplo, polímeros de dois tipos de bases orgânicas como uracila e citosina (poli-UC) foram combinados para estudos do código genético.

**Corpos embrioides.** Massas de células diferenciadas e indiferenciadas derivadas de células-tronco embrionárias.

**Corpos polares.** Em animais fêmeas, as células menores produzidas na meiose que não se transformam em ovócitos. O primeiro corpo polar é produzido na divisão I e não pode passar pela divisão II. O segundo corpo polar é produzido na divisão II.

**Corpúsculo basal.** Grânulo pequeno ao qual está fixado um cílio ou flagelo.

**Corpúsculo de Barr.** Massa condensada de cromatina que é encontrada nos núcleos de mamíferos placentários e que contém um ou mais cromossomos X; nome dado em homenagem ao seu descobridor, Murray Barr.



- Correlação.** Associação estatística entre variáveis.
- Correpressor.** Molécula efetora que forma um complexo com um repressor e desativa a expressão de um gene ou conjunto de genes.
- Cosmídios.** Vetores de clonagem que são híbridos entre cromossomos de fago  $\lambda$  e plasmídios; contêm sítios *cos* do fago  $\lambda$  e origens de replicação dos plasmídios.
- Covariância.** Medida da associação estatística entre variáveis.
- cpDNA.** DNA de plastídios vegetais, entre eles os cloroplastos.
- Cromátide.** Na mitose ou meiose, um dos dois filamentos idênticos resultantes da autoduplicação de um cromossomo.
- Cromátide-irmã.** Um dos produtos da duplicação do cromossomo.
- Cromatina.** O complexo de DNA e proteínas em cromossomos eucarióticos; originalmente nomeada em razão da rapidez com que é tingida por alguns corantes.
- Cromatografia.** Método para separar e identificar os componentes de misturas de moléculas com propriedades químicas e físicas semelhantes.
- Cromocentro.** Corpo produzido por fusão das regiões heterocromáticas dos cromossomos nos tecidos politênicos (p. ex., glândulas salivares) de alguns dípteros.
- Cromômeros.** Corpos pequenos identificados por seu tamanho característico e arranjo linear ao longo de um cromossomo.
- Cromonema.** Filamento distinguido opticamente que forma uma estrutura axial dentro de cada cromossomo.
- Cromossomo acêntrico.** Fragmento de cromossomo sem centrômero.
- Cromossomo balanceador.** Em genética de *Drosophila*, cromossomo com múltiplas inversões e marcador dominante que inibe a recombinação com um cromossomo homólogo estruturalmente normal.
- Cromossomo CIB.** Cromossomo X de *Drosophila* que tem uma mutação causadora de olhos em forma de barra e uma mutação letal recessiva dentro de uma grande inversão.
- Cromossomo composto.** Cromossomo formado pela união de dois cromossomos distintos do mesmo par, como em cromossomos X ligados ou cromossomos X-Y ligados.
- Cromossomo dicêntrico.** Cromossomo que tem dois centrômeros.
- Cromossomo heterólogo.** Cromossomo que contém um conjunto de genes diferente do cromossomo ao qual é comparado.
- Cromossomo metacêntrico.** Cromossomo que tem o centrômero perto do meio e dois braços de comprimentos aproximadamente iguais.
- Cromossomo X.** Cromossomo associado à determinação do sexo. Na maioria dos animais, a fêmea tem dois cromossomos X e o macho, um.
- Cromossomo Y.** O par do cromossomo X no macho de muitas espécies de animais.
- Cromossomos.** Aglomerados de nucleoproteína de coloração escura que são observados nas células durante a divisão. Cada cromossomo tem um arranjo linear de genes.
- Cromossomos homólogos.** Cromossomos que ocorrem em pares e geralmente têm tamanho e formato semelhantes, um proveniente do genitor de sexo masculino e o outro, do genitor de sexo feminino. Esses cromossomos contêm o mesmo arranjo de genes.
- Cromossomos politênicos.** Cromossomos gigantes produzidos por replicação na intérfase sem divisão e constituídos de muitas cromátides idênticas dispostas lado a lado, em padrão semelhante a um cabo.
- Cromossomos sexuais.** Cromossomos associados à determinação do sexo.
- Crossing over desigual.** *Crossing over* entre sequências de DNA repetidas com pareamento desalinhado, criando produtos duplicados e deficientes.
- Crossing over.** Processo no qual os cromossomos trocam material por meio da quebra e reunião de suas moléculas de DNA. (Veja também **Recombinação**.)
- Cruzamento consanguíneo.** Cruzamento entre parentes.
- Cruzamento de irmãos.** Cruzamentos de dois indivíduos que têm os mesmos pais; cruzamento entre irmão e irmã.
- Cruzamento inter-racial ou interespecífico.** Cruzamento de membros pertencentes a raças ou espécies diferentes.
- Cruzamento mono-híbrido.** Cruzamento entre indivíduos que diferem em apenas uma característica ou nos quais só está sendo considerado um traço.
- Cruzamento preferencial.** Cruzamento no qual os parceiros são escolhidos em razão da semelhança fenotípica.
- Cruzamentos recíprocos.** Cruzamentos entre diferentes linhagens com os sexos invertidos; por exemplo, fêmea A  $\times$  macho B e macho A  $\times$  fêmea B são cruzamentos recíprocos.
- Cruzamento-teste.** Retrocruzamento do tipo parental recessivo, ou cruzamento entre indivíduos de genótipo desconhecido com um testador totalmente recessivo para determinar se o indivíduo em questão é heterozigoto ou homozigoto para determinado alelo. Também é usado como teste de ligação gênica.

## D

- Dáton.** A massa de um átomo de hidrogênio.
- de novo.** Original, novo, mais uma vez.
- Deficiência (deleção).** Ausência de um segmento de um cromossomo, reduzindo o número de *loci*.
- Degeneração (do código genético).** Especificação de um aminoácido por mais de um códon.
- Depressão endogâmica.** Observação de que as linhagens endogâmicas são mais fracas que as linhagens não endogâmicas.
- Deriva.** Veja **Deriva genética aleatória**.
- Deriva genética.** Veja **Deriva genética aleatória**.
- Deriva genética aleatória.** Alterações da frequência alélica em pequenas populações reprodutivas decorrentes de variações fortuitas.
- Desnaturação.** Perda da configuração nativa de uma macromolécula, geralmente acompanhada de perda da atividade biológica. As proteínas desnaturadas costumam desdobrar suas cadeias polipeptídicas e exibir propriedades alteradas de solubilidade.
- Desoxirribonuclease (DNase).** Enzima que hidrolisa o DNA.
- Desrepressão.** Processo de ativar a expressão de um gene ou conjunto de genes que havia sido reprimida (desativada).
- Desvio.** Em estatística, afastamento de um valor esperado.
- Desvio padrão.** Medida de variabilidade em um conjunto de dados; a raiz quadrada da variância.



- Determinação.** Processo de comprometimento das células indiferenciadas de um embrião para transformação em tipos celulares específicos, como neurônio, fibroblasto e célula muscular.
- Diacinese.** Estágio da meiose logo antes da metáfase I no qual há encurtamento e espessamento dos bivalentes.
- Dicotiledônea.** Vegetal com dois cotilédones (folhas embrionárias).
- Diferenciação.** Processo em que células não especializadas desenvolvem estruturas e funções características.
- Diferencial de seleção.** Na reprodução de vegetais e animais, a diferença entre a média dos indivíduos selecionados para serem pais e a média da população em geral.
- Di-híbrido, cruzamento di-híbrido.** Indivíduo heterozigoto para dois pares de alelos; a prole de um cruzamento entre homozigotos que difere em dois aspectos.
- Dímero.** Substância que tem composição percentual igual à de outra, mas o dobro do peso molecular; produzido por polimerização.
- Dimorfismo.** Duas formas diferentes em um grupo em relação a características como sexo, tamanho ou coloração.
- Diploide.** Organismo ou célula com dois conjuntos de cromossomos ( $2n$ ) ou dois genomas. Os tecidos somáticos de vegetais superiores e animais geralmente têm constituição cromossômica diploide em contraste com os gametas haploides (monoploides).
- Diplóteno (diplonema).** Estágio na prófase da meiose I depois do paquíteno, mas antes da diacinese, no qual os cromossomos de bivalentes se separam nos centrômeros e na região adjacente.
- Disco imaginal.** Massa de células nas larvas de *Drosophila* e de outros insetos holometabólicos que dá origem a determinado órgão no adulto, como antena, olho ou asa.
- Discordante.** Membros de um par que têm características diferentes, em vez de semelhantes.
- Disgenesia híbrida.** Em *Drosophila*, uma síndrome de características anormais da linhagem germinativa, que incluem mutação, quebra do cromossomo e esterilidade, resultantes da atividade do elemento transponível.
- Disjunção.** Separação de cromossomos homólogos durante a anáfase de divisões mitóticas ou meióticas. (Veja também **Não disjunção**).
- Dissociação (Ds).** Elemento transponível no milho, originalmente detectado como agente mediador da quebra do cromossomo em resposta ao efeito do *Activator (Ac)*, outro elemento transponível.
- Distribuição de frequência.** Gráfico que mostra a incidência relativa ou absoluta de classes em uma população. As classes podem ser definidas por uma variável distinta ou contínua; neste último caso, cada classe representa um intervalo diferente na escala de medida.
- Distribuição independente.** Distribuição aleatória de alelos para os gametas que ocorre quando os genes estão localizados em diferentes cromossomos. A distribuição de um par de alelos é independente de outros genes localizados em cromossomos não homólogos.
- Divisão equacional.** Divisão do tipo mitótica que geralmente é a segunda divisão na sequência meiótica; mitose somática e divisão não reducional da meiose.
- Divisão reducional.** Fase da meiose na qual há separação dos cromossomos maternos e paternos do bivalente (cf. **Divisão equacional**).
- DNA.** Ácido desoxirribonucleico; material genético que contém informações e constitui os genes. O DNA é uma macromolécula composta de uma cadeia longa de desoxirribonucleotídeos unidos por ligações fosfodiéster. Cada desoxirribonucleotídeo contém um grupo fosfato, o açúcar 2-desoxirribose com cinco carbonos e uma base nitrogenada.
- DNA do cloroplasto.** Veja **cpDNA**.
- DNA fotoliase.** Enzima que usa energia da luz azul para clivar ligações cruzadas covalentes induzidas por luz ultravioleta na timina, na citosina e nos dímeros citosina-timina no DNA.
- DNA girase.** Enzima bacteriana que catalisa a super-helicoidização negativa do DNA.
- DNA helicase.** Enzima que catalisa o desenrolamento dos filamentos complementares de uma dupla hélice de DNA.
- DNA ligase.** Enzima que catalisa o fechamento covalente de cortes nas duplas hélices de DNA.
- DNA mitocondrial.** Veja **mtDNA**.
- DNA polimerase.** Enzima que catalisa a síntese de DNA.
- DNA primase.** Enzima que catalisa a síntese de filamentos curtos de RNA que iniciam a síntese de filamentos de DNA.
- DNA repetitivo.** Sequências de DNA presentes em múltiplas cópias em um genoma – em algumas ocasiões, um milhão de vezes ou mais.
- DNA satélite.** Componente do genoma que pode ser isolado do restante do DNA por centrifugação em gradiente de densidade. Em geral, é constituído de sequências curtas e altamente repetitivas.
- DNA topoisomerase.** Enzima que catalisa a introdução ou a retirada de super-hélices do DNA.
- Doença de Huntington (DH).** Distúrbio neurodegenerativo de início tardio (30 a 50 anos de idade) em seres humanos, causado por mutação autossômica dominante. O defeito genético é uma repetição expandida do trinucleotídeo (CAG)<sub>n</sub> que codifica uma região poliglutamina anormalmente longa perto da terminação amino do produto do gene *huntingtina*.
- Doenças autoimunes.** Distúrbios nos quais o sistema imune dos indivíduos afetados produz anticorpos contra antígenos próprios – antígenos sintetizados pelas próprias células.
- Dominância incompleta.** Expressão de dois alelos em um heterozigoto que permite distinguir o heterozigoto dos pais homozigotos.
- Dominância influenciada pelo sexo.** Expressão dominante que depende do sexo do indivíduo. Por exemplo, em algumas raças de carneiros, os chifres são dominantes nos machos e recessivos nas fêmeas.
- Dominante.** Termo usado para se referir a um alelo que se manifesta para a exclusão de um alelo diferente em um organismo heterozigoto.
- Domínio aptâmero.** A região de um riborregulador (*riboswitch*) que se liga ao metabólito.
- Domínio de expressão.** Região de um riborregulador (*riboswitch*) que pode se dobrar de duas formas, uma que facilita a expressão gênica e outra que bloqueia a expressão gênica.



**Down, síndrome de.** Fenótipo resultante de trissomia do cromossomo 21 em seres humanos.

**Dupla hélice.** Molécula de DNA composta de dois filamentos complementares.

**Duplicação.** Ocorrência de um segmento mais de uma vez no mesmo cromossomo ou genoma; também se refere à multiplicação de células.

**Duplicação do sítio-alvo.** Sequência de DNA duplicada quando há inserção de um elemento transponível; geralmente encontrada em cada extremidade da inserção.

## E

**Ecdisona.** Hormônio que influencia o desenvolvimento em insetos.

**Eclosão.** Saída de um inseto adulto do estágio pupal.

**Ecótipo.** População ou linhagem de organismos adaptados a determinado *habitat*.

**Ectópico.** Termo usado para descrever um fenômeno ocorrido em local anormal.

**Edição de RNA.** Processos pós-transcrição que alteram as informações codificadas em transcritos de genes (RNA).

**Efeito de posição *cis-trans*.** Ocorrência de diferentes fenótipos quando há duas mutações em heterozigotos *cis* e *trans*.

**Efeito materno.** Característica controlada por um gene da mãe, mas expresso na prole.

**Efeitos alélicos aditivos.** Fatores genéticos que elevam ou reduzem o valor de um fenótipo em uma escala linear de medida.

**Elemento controlador.** No milho, um elemento transponível como *Ac* ou *Ds* capaz de influenciar a expressão de um gene adjacente.

**Elemento genético transponível.** Elemento de DNA capaz de passar de um local para outro no genoma.

**Elemento IS (sequência de inserção).** Sequência de DNA curta (800 a 1.400 pares de nucleotídeos) encontrada em bactérias e capaz de se transpor para uma nova localização genômica; outras sequências de DNA limitadas por elementos IS também podem ser transpostas.

**Elemento P.** Elemento transponível em *Drosophila* que, quando ativado, causa disgenesia híbrida.

**Elemento semelhante a retrovírus.** Tipo de retrotranspósom que se assemelha à forma integrada de um retrovírus.

**Eletroforese.** Migração de partículas suspensas em um campo elétrico.

**Eletroporação.** Processo que torna as membranas celulares permeáveis ao DNA mediante aplicação de corrente elétrica intensa.

**Embrião.** Organismo nos estágios iniciais do desenvolvimento; em seres humanos, os primeiros 2 meses de vida intrauterina.

**Endogamia.** Cruzamento entre indivíduos com parentesco.

**Endomitose.** Duplicação de cromossomos sem divisão do núcleo, com consequente aumento do número de cromossomos na célula. Os filamentos dos cromossomos se separam, mas a célula não se divide.

**Endonuclease de restrição.** Veja **Enzima de restrição**.

**Endonuclease.** Enzima que quebra filamentos de DNA em posições internas; algumas participam da recombinação do DNA.

**Endopoliploidia.** Estado em que as células de um organismo diploide contêm múltiplos do número diploide de cromossomos (ou seja,  $4n$ ,  $8n$ , e assim por diante).

**Endosperma.** Tecido nutritivo que se desenvolve no saco embrionário da maioria das angiospermas. De modo geral, surge depois da fertilização dos dois núcleos endospermícos primários fundidos do saco embrionário por um dos dois núcleos espermáticos. Na maioria dos vegetais diploides, o endosperma é triploide ( $3n$ ).

**Endossimbiose.** Relação mutuamente benéfica na qual um organismo vive dentro de outro.

**Enzima.** Proteína que acelera uma reação química específica em um sistema vivo.

**Enzima constitutiva.** Enzima sintetizada continuamente quaisquer que sejam as condições de crescimento (cf. **Enzima induzível** e **Enzima repressível**).

**Enzima de restrição.** Endonuclease que reconhece uma sequência curta específica no DNA e cliva a molécula de DNA naquele sítio ou perto dele.

**Enzima induzível.** Enzima sintetizada apenas na presença do substrato que atua como indutor.

**Enzima repressível.** Enzima cuja síntese é diminuída por uma molécula reguladora.

**Enzimas de reparo do DNA.** Enzimas que catalisam o reparo do DNA danificado.

**Epigenético.** Termo relativo às causas não genéticas de um fenótipo.

**Epissomo.** Elemento genético que pode estar presente ou ausente em diferentes células e que pode estar inserido em um cromossomo ou independente no citoplasma (p. ex., o fator de fertilidade [F] em *Escherichia coli*).

**Epistasia.** Interações entre produtos de genes não alélicos. Os genes suprimidos são denominados hipostáticos. A dominância está associada a membros de pares alélicos, ao passo que a epistasia resulta de interações dos produtos de não alelos.

**Equilíbrio.** Estado de sistemas dinâmicos em que não há alteração real.

**Equilíbrio da ligação.** Estado no qual há associação aleatória entre os alelos de *loci* ligados nos cromossomos de uma população.

**Equilíbrio genético.** Condição em um grupo de organismos em que há intercruzamento na qual as frequências de alelos permanecem constantes ao longo do tempo.

**Erro padrão.** Medida da variação em uma série de médias.

**Especiação alopátrica.** Especiação decorrente, ao menos em parte, do isolamento geográfico.

**Especiação simpátrica.** Formação de uma nova espécie por populações que habitam as mesmas regiões geográficas ou regiões superpostas.

**Espécie.** Intercruzamento, populações naturais que são isoladas reprodutivamente de outros grupos do mesmo tipo.

**Espermátides.** As quatro células formadas pelas divisões meióticas na espermatogênese. As espermátides tornam-se espermatozoides maduros.

**Espermatócito (célula-mãe do espermatozoide).** A célula que passa por duas divisões meióticas (espermatogênese) para formar quatro espermátides; o espermatócito *primário* antes da conclusão da primeira divisão meiótica; o esperma-



tócito *secundário* depois da conclusão da primeira divisão meiótica.

**Espematócito secundário.** Veja **Espematócito**.

**Espematogênese.** Processo de maturação dos gametas (espermatozoides) masculinos.

**Espematogônia.** Célula germinativa masculina primordial que pode se dividir por mitose e produzir mais espermatogônias. A espermatogônia pode entrar em fase de crescimento e dar origem a um espermátócito primário.

**Espermatozoide.** Célula germinativa masculina madura.

**Espermiogênese.** Formação de espermatozoides a partir das espermátides; a parte da espermatogênese que sucede as divisões meióticas dos espermátócitos.

**Espliceossomo.** O complexo de RNA/proteína que excisa íntrons de transcritos primários de genes nucleares em eucariotos.

**Esporófito.** Geração diploide no ciclo de vida de um vegetal que produz esporos haploides por meiose.

**Esqueleto (*scaffold*).** A estrutura central de cromossomos condensados. É constituído de proteínas cromossômicas não histônicas.

**EST (etiquetas de sequências expressas).** Sequências curtas de cDNA usadas para ligar mapas físicos e mapas genéticos (RFLP).

**Estame.** Estrutura alongada que sustenta as anteras em angiospermas.

**Estatística.** Valor baseado em uma amostra ou amostras de uma população a partir do qual é possível fazer estimativas de um valor ou parâmetro populacional.

**Esterilidade.** Incapacidade de gerar prole.

**Estrogênio.** Hormônio feminino ou substância produtora de estro.

**Eucarioto.** Membro do grande grupo de organismos que têm núcleos celulares delimitados por membrana (cf. **Procarioto**).

**Eucromatina.** Material genético que não é tingido tão intensamente por alguns corantes durante a interfase e que constitui muitos tipos diferentes de genes (cf. **Heterocromatina**).

**Eugenia.** Aplicação dos princípios de Genética à melhora da humanidade.

**Euploide.** Organismo ou célula cujo número de cromossomo é um múltiplo exato do número monoploide ( $n$ ) ou haploide. Os termos usados para identificar diferentes níveis em uma série euploide são diploide, triploide, tetraploide, e assim por diante (cf. **Aneuploide**).

**Excinuclease.** Complexo proteico contendo endonuclease que excisa um segmento de DNA lesado durante o reparo por excisão.

**Exogamia.** Cruzamento de indivíduos sem parentesco.

**Éxons.** Os segmentos de um gene eucariótico que correspondem às sequências no transcrito de RNA processado final desse gene.

**Exonuclease.** Enzima que digere DNA ou RNA, partindo das extremidades dos filamentos.

**Expansão binomial.** Multiplicação exponencial de uma expressão constituída de dois termos unidos por um sinal de adição (+) ou subtração (−), como  $(a + b)^n$ .

**Expressão gênica.** O processo pelo qual genes produzem RNA e proteínas e exercem seus efeitos sobre o fenótipo de um organismo.

**Extracromossômico.** Designa estruturas que não são parte dos cromossomos; unidades de DNA no citoplasma que controlam a herança citoplasmática.

## F

**F<sub>1</sub>.** A primeira geração filial; a primeira geração de descendentes de determinado cruzamento.

**F<sub>2</sub>.** A segunda geração filial produzida por cruzamento *inter se* ou por autopolinização de F<sub>1</sub>. Os "netos" endogâmicos de determinado cruzamento, mas em experimentos genéticos controlados; a autofertilização de F<sub>1</sub> (ou equivalente) está implícita.

**Fago.** Veja **Bacteriófago**.

**Fago lítico.** Veja **Fago virulento**.

**Fago temperado.** Fago (vírus) que invade, mas não destrói (lisa) o hospedeiro (célula bacteriana) (cf. **Fago virulento**). No entanto, pode entrar em seguida no ciclo lítico.

**Fago virulento.** Fago (vírus) que destrói a célula hospedeira (bacteriana) (cf. **Fago temperado**).

**Fagomídios.** Vetores de clonagem que contêm componentes derivados tanto dos plasmídios quanto dos cromossomos dos fagos.

**Família multigênica.** Grupo de genes que têm sequência nucleotídica semelhante ou que produzem polipeptídios com sequências de aminoácidos semelhantes.

**Fase de ligação gênica.** Arranjo de marcadores genéticos ligados em um heterozigoto. Os marcadores podem estar na fase de acoplamento ( $A B/a b$ ) ou de repulsão ( $A b/a B$ ).

**Fator de resistência.** Plasmídio que confere a uma bactéria resistência contra antibióticos.

**Fator de transcrição.** Proteína que regula a transcrição de genes.

**Fator determinante testicular (TDF).** Proteína produzida no início do desenvolvimento de mamíferos do sexo masculino que estimula a diferenciação dos testículos a partir das gônadas embrionárias.

**Fator F.** Epissomo bacteriano que confere a capacidade de atuar como doador genético ("masculino") na conjugação; o fator de fertilidade em bactérias.

**Fator sexual.** Epissomo bacteriano (p. ex., plasmídio F em *E. coli*) que possibilita que a célula seja doadora de material genético. O fator sexual pode ser propagado no citoplasma ou integrado ao cromossomo bacteriano.

**Fator sigma.** Subunidade de RNA polimerases procarióticas responsável pelo início da transcrição em sequências de iniciação específicas.

**Fatores de alongamento.** Proteínas solúveis necessárias para o alongamento da cadeia polipeptídica.

**Fatores de iniciação.** Proteínas solúveis necessárias para o início da tradução.

**Fatores de liberação (RF).** Proteínas solúveis que reconhecem códons de término em mRNA e terminam a tradução em resposta a esses códons.

**Fatores de transcrição basais.** Proteínas necessárias para o início da transcrição em eucariotos.

**Fenilalanina.** Veja **Aminoácido**.

**Fenilcetonúria.** Distúrbio metabólico que acarreta retardo mental; com herança recessiva mendeliana e tratado no início da infância por meio de dieta especial.



**Fenocópia.** Organismo cujo fenótipo (mas não genótipo) foi modificado pelo ambiente para se assemelhar ao fenótipo de um organismo diferente (mutante).

**Fenótipo.** Características observáveis de um organismo.

**Fertilização.** Fusão de um gameta masculino (espermatozoide ou anterozoide) com um gameta feminino (ovócito ou oosfera) para formar um zigoto.

**Feto.** Estágio pré-natal de um animal vivíparo entre o estágio embrionário e o nascimento; em seres humanos, os últimos 7 meses antes do nascimento.

**Fibra de cromatina.** Unidade básica da organização de cromossomos eucarióticos, constituída de DNA e proteínas associadas montados em um filamento com diâmetro médio de 30 nm.

**Fibrose cística (FC).** Distúrbio autossômico recessivo em seres humanos caracterizado por obstrução dos pulmões, do pâncreas e do fígado por muco e, consequentemente, por infecções crônicas. A expectativa de vida média de um indivíduo com fibrose cística é de 35 anos.

**Filamento contínuo (líder).** Filamento de DNA sintetizado de maneira contínua durante a replicação.

**Filamento descontínuo (atrasado).** Filamento de DNA sintetizado de maneira descontínua durante a replicação.

**Filamento não molde (codificante).** Na transcrição, o filamento não transcrito de DNA. A sequência é igual à do transcrito de RNA, exceto pela presença de T nas posições em que há U no transcrito de RNA.

**Filamento sense (de RNA).** Veja *RNA sense*.

**Filamento-molde.** Na transcrição, o filamento de DNA que é copiado para produzir um filamento complementar de RNA.

**Filial.** Veja  $F_1$  e  $F_2$ .

**Filogenia.** Diagrama que mostra as relações evolutivas entre um grupo de organismos; árvore evolutiva.

**Fissão.** Mecanismo de divisão celular entre os procariotos nos quais o material genético da célula-mãe é duplicado e, depois, dividido igualmente pelas duas células-filhas.

**Fixação.** Evento que ocorre quando todos os alelos em um *locus*, exceto um, são eliminados de uma população. Diz-se que o alelo remanescente, com frequência de 100%, foi fixado.

**Flagelo (adj. flagelado).** Organela de locomoção, semelhante a um chicote, existente em determinadas células; estruturas locomotoras em protozoários flagelados.

**Fluxo gênico.** A propagação de genes de uma população reprodutiva para outra por migração, possivelmente levando a variações da frequência do alelo.

**Forquilha de replicação.** Estrutura em formato de Y em que os dois filamentos parentais de uma dupla hélice de DNA são desenrolados e usados como moldes para a síntese de novos filamentos complementares.

**Fotorreativação.** Processo de reparo do DNA fotodependente.

**Fragmento de restrição.** Fragmento de DNA produzido por clivagem de uma molécula de DNA com uma ou mais endonucleases de restrição.

**Frequência alélica.** A proporção de um alelo em relação a todos os alelos de um *locus* em uma população.

**Fuso.** Sistema de microtúbulos que distribui cromossomos duplicados de modo igual e exato para cada célula-filha de uma célula eucariótica que se divide.

## G

**Galha.** Tumor dos vegetais.

**Gameta.** Célula reprodutiva masculina ou feminina madura (espermatozoide/anterozoide ou ovócito/oosfera).

**Gametófito.** Fase do ciclo de vida dos vegetais em que há produção de gametas; as células têm  $n$  cromossomos.

**Gametófito feminino.** Espaço grande de paredes finas dentro do óvulo das espermatófitas que contém os oito núcleos haploides idênticos derivados da mitose do megasporo produzido por meiose.

**Gametófito masculino.** Os três núcleos haploides idênticos em um grão de pólen.

**Gametogênese.** Formação de gametas.

**Gástrula.** Embrião animal inicial constituído de duas camadas de células; um estágio embriológico depois da blástula.

**Gêmeos dizigotos (DZ).** Gêmeos de dois zigotos ou fraternos.

**Gêmeos monozigóticos.** Gêmeos originados de um único zigoto ou idênticos.

**GenBank.** Banco de dados de sequências de DNA mantido pelo National Center for Biotechnology Information, no National Institutes of Health nos EUA. Existem bancos de dados semelhantes na Europa (European Molecular Biology Laboratory Data Library) e no Japão (DNA DataBank of Japan).

**Gene.** Determinante hereditário de uma função biológica específica; unidade de herança (DNA) localizada em posição fixa no cromossomo; segmento de DNA que codifica um polipeptídeo e cuja definição operacional é dada pelo teste *cis-trans* ou de complementação.

**Gene constitutivo.** Gene continuamente expresso em todas as células de um organismo.

**Gene de ancoragem.** Gene posicionado tanto no mapa físico quanto no mapa genético de um cromossomo.

**Gene de efeito materno.** Gene cujo produto atua na prole do indivíduo de sexo feminino que tem o gene.

**Gene de pair-rule (gene de paridade segmentar).** Gene que influencia a formação de segmentos do corpo de *Drosophila*.

**Gene estrutural.** Gene que especifica a síntese de um polipeptídeo.

**Gene gap.** Gene que controla a formação de segmentos adjacentes no corpo da *Drosophila*.

**Gene induzível.** Gene expresso apenas na presença de um metabólito específico, o indutor.

**Gene marcador selecionável dominante.** Gene que permite a sobrevivência da célula hospedeira em condições em que, de outro modo, morreria.

**Gene marcador selecionável quimérico.** Gene construído com sequências de DNA de duas ou mais origens que permite a sobrevivência de uma célula ou organismo em condições em que normalmente morreria.

**Gene pseudoautossômico.** Gene localizado tanto no cromossomo X quanto no cromossomo Y.

**Gene regulador.** Gene que controla a taxa de expressão de outro gene ou genes. Exemplo: o gene *lacI* produz uma proteína que controla a expressão dos genes estruturais do operon *lac* em *Escherichia coli*.

**Gene repressor.** Gene que codifica um repressor.



**Gene seletor.** Gene que influencia o desenvolvimento de segmentos específicos do corpo em *Drosophila*; um gene homeótico.

**Gene supressor tumoral.** Gene cujo produto participa da repressão à divisão celular.

**Genes de polaridade de segmento.** Grupo de genes cujos produtos definem os compartimentos anterior e posterior de cada segmento que se forma ao longo do eixo anteroposterior de embriões de *Drosophila*.

**Genes de segmentação.** Grupo de genes que controlam o desenvolvimento inicial de embriões de *Drosophila*. Seus produtos definem segmentos ao longo do eixo anteroposterior.

**Genes homeóticos.** Grupo de genes cujos produtos controlam a formação do corpo de um embrião por controle da expressão de outros genes em regiões segmentares ao longo do eixo anteroposterior.

**Genes homólogos.** Genes que evoluíram a partir de um gene ancestral comum (cf. **Genes ortólogos**; **Genes parálogos**).

**Genes mutáveis.** Genes cuja taxa de mutação é muito alta.

**Genes ortólogos.** Genes homólogos presentes em diferentes espécies (cf. **Genes homólogos**).

**Genes parálogos.** Genes homólogos presentes em uma espécie (cf. **Genes homólogos**).

**Genética.** A ciência da hereditariedade e da variação.

**Genética de populações.** Ramo da genética que estuda as frequências de alelos e genótipos nas populações reprodutivas.

**Genética reversa.** Técnicas genéticas que usam a sequência nucleotídica de um gene a fim de criar procedimentos para isolamento de mutações no gene ou para bloqueio de sua expressão.

**Genoma dobrado.** O estado intracelular condensado do DNA no nucleóide de uma bactéria. O DNA é segregado em domínios, e há super-helicoidização negativa independente de cada domínio.

**Genoma.** Conjunto completo ( $n$ ) de cromossomos (portanto, de genes) herdado como uma unidade de um dos pais.

**Genômica comparativa.** Ramo da genômica que compara a estrutura e a função dos genomas de diferentes espécies.

**Genômica.** Estudo da estrutura e da função de genomas inteiros.

**Genótipo.** Constituição genética (composição gênica) de um organismo (cf. **Fenótipo**).

**Germoplasma.** Material hereditário transmitido à prole por intermédio das células germinativas.

**Ginandromorfo.** Indivíduo que tem uma parte do corpo feminina e a outra, masculina; um mosaico sexual.

**Glicocorticoide.** Hormônio esteroide que regula a expressão gênica em animais superiores.

**Globulinas.** Proteínas comuns do sangue que são insolúveis em água e solúveis em soluções salinas. É possível distinguir as alfa, beta e gamaglobulinas no soro humano. As gamaglobulinas são importantes no desenvolvimento da imunidade a doenças.

**Gônada.** Órgão sexual (*i. e.*, ovário ou testículo) que produz gametas.

**Grão de pólen.** Gametófito masculino nos vegetais superiores.

**Graus de liberdade.** Índice associado à distribuição de frequência de uma estatística de teste calculada a partir de dados amostrais.

**Guanina (G).** Base purínica encontrada no DNA e no RNA.

## H

**Haploide (monoploide).** Organismo ou célula que tem apenas um conjunto completo ( $n$ ) de cromossomos ou um genoma.

**Haplótipo.** Conjunto de variantes genéticas ligadas, principalmente por polimorfismos de nucleotídeo único (SNP), em um cromossomo.

**Haptoglobina.** Proteína sérica, alfa globulina, presente no sangue.

**Hélice.** Qualquer estrutura que tenha formato espiral. O modelo de Watson e Crick de DNA tem a forma de uma dupla hélice.

**Hemizigoto.** Indivíduo que tem uma cópia de um cromossomo ou gene, como na ligação ao sexo ou em consequência da deleção.

**Hemofilia.** Doença hemorrágica; tendência ao sangramento excessivo mesmo em caso de pequeno ferimento; distúrbio hereditário dependente de um gene recessivo ligado ao sexo.

**Hemoglobina.** Proteína conjugada composta que contém ferro, localizado em eritrócitos de vertebrados; importante no transporte de oxigênio para as células do corpo.

**Hemolinfa.** Mistura de sangue e outros líquidos na cavidade corporal de um invertebrado.

**Herança citoplasmática.** Transmissão hereditária dependente do citoplasma ou de estruturas no citoplasma, não dos genes nucleares; herança extracromossômica. Exemplo: as características dos plastídios nos vegetais podem ser herdadas por um mecanismo independente dos genes nucleares.

**Herança materna.** Herança controlada por fatores extracromossômicos (*i. e.*, citoplasmáticos) transmitidos pelo gameta feminino.

**Herança quantitativa.** Herança de características mensuráveis (altura, peso, intensidade da cor) que dependem da ação acumulativa de muitos genes, cada um deles produzindo um pequeno efeito no fenótipo.

**Herdabilidade.** Grau de controle de uma característica pela herança. (Ver também **Herdabilidade em sentido amplo** e **Herdabilidade em sentido restrito**).

**Herdabilidade em sentido amplo.** Em genética quantitativa, a proporção da variância fenotípica total devida a fatores genéticos.

**Herdabilidade em sentido restrito.** Em genética quantitativa, a proporção da variância fenotípica devida à soma dos efeitos dos alelos.

**Hereditariedade.** Semelhança entre indivíduos associada à descendência; transmissão de características dos pais para os filhos.

**Heredograma.** Tabela, quadro ou diagrama que representa a ascendência de um indivíduo.

**Hermafrodita.** Indivíduo que tem órgãos reprodutivos masculinos e femininos.

**Heteroalelos.** Mutações que são alélicas do ponto de vista funcional, mas não alélicas do ponto de vista estrutural; mutações em diferentes sítios de um gene.

**Heterocromatina.** Cromatina que adquire coloração escura mesmo durante a interfase, muitas vezes contém DNA repetitivo com poucos genes.

**Heterodúplex.** Ácido nucleico bifilamentar que contém um ou mais pares de bases incompatíveis (não complementares).



**Heterose.** Superioridade de genótipos heterozigotos em relação a uma ou mais características em comparação com os homozigotos correspondentes.

**Heterozigosidade.** Proporção de indivíduos heterozigotos em uma população; usada como medida de variabilidade genética.

**Heterozigoto.** Organismo com membros diferentes de determinado par ou série de alelos que, conseqüentemente, produz gametas diferentes.

**Heterozigoto *cis*.** Heterozigoto que apresentam duas mutações em configuração *cis* – por exemplo,  $a^+ b^+ / a b$ .

**Heterozigoto *trans*.** Heterozigoto que contém duas mutações em configuração *trans* – por exemplo,  $a b^+ / a^+ b$ .

**Hfr.** Linhagem de alta frequência de recombinação de *Escherichia coli*; nessas linhagens, o episômo F é integrado ao cromossomo bacteriano.

**Hibridização de colônia ou placa *in situ*.** Procedimento para exame de colônias ou placas que crescem em membranas ou placas de cultura a fim de pesquisar sequências de DNA específicas pela hibridização de sondas de ácidos nucleicos com as moléculas de DNA presentes nessas colônias ou placas.

**Hibridização.** Intercruzamento de espécies, raças, variedades, e assim por diante, de vegetais ou animais; processo de formação de um híbrido por polinização cruzada de vegetais ou por acasalamento de animais de diferentes tipos.

**Hibridização *in situ*.** Método para determinar a localização de sequências de DNA específicas em cromossomos, por hibridização de DNA ou RNA marcado com DNA desnaturado em preparações cromossômicas, e visualizar a sonda hibridizada por autorradiografia ou microscopia de fluorescência.

**Hibridização *in situ* com fluorescência (FISH).** Hibridização *in situ* que usa uma sonda de DNA ou RNA ligada a um corante fluorescente.

**Híbrido.** Prole de pais homozigotos que difere em um ou mais genes; em um sentido amplo, a prole de um cruzamento de linhagens sem parentesco.

**Hipermutação somática.** Alta frequência de mutação ocorrida nos segmentos gênicos que codificam as regiões variáveis de anticorpos durante a diferenciação de linfócitos B em plasmócitos produtores de anticorpos.

**Hiperploide.** Condição genética na qual há um número excessivo de um cromossomo ou de um segmento de cromossomo no genótipo (cf. **Hipoploide**).

**Hipomórfico.** Termo designativo de um alelo mutante que tem menor expressão que um alelo do tipo selvagem, mas que não suprime totalmente a expressão. Esse alelo mutante é denominado hipomorfo.

**Hipoploide.** Condição genética na qual há um número menor que o normal de um cromossomo ou de um segmento de cromossomo no genótipo (cf. **Hiperploide**).

**Hipótese.** Em ciência, proposição para explicar um fenômeno.

**Hipótese da oscilação.** Hipótese que explica como um tRNA reconhece dois códons. Há pareamento correto das duas primeiras bases do códon e anticódon de mRNA, mas a terceira base no anticódon apresenta movimento (oscilação) que permite o pareamento com mais de uma base.

**Hipótese dos fatores múltiplos.** Teoria proposta por R. A. Fisher e outros para explicar a variação de fenótipos complexos como altura, peso e suscetibilidade a doenças.

**Histonas.** Grupo de proteínas ricas em aminoácidos básicos. Atuam no espiralamento de DNA em cromossomos e na regulação da atividade gênica.

**HIV (vírus da imunodeficiência humana).** Retrovírus causador da AIDS em seres humanos.

**Holoenzima.** Forma de uma enzima multimérica na qual estão presentes todos os polipeptídios componentes.

**Homeobox.** Sequência de DNA encontrada em vários genes participantes da especificação de órgãos em diferentes partes do corpo em animais; característica de genes que influenciam a segmentação em animais. O homeobox corresponde a uma sequência de aminoácidos no polipeptídeo codificado por esses genes; essa sequência é denominada homeodomínio.

**Homeodomínio.** Veja **Homeobox**.

**Homoalelos.** Mutações alélicas tanto do ponto de vista funcional quanto estrutural; mutações no mesmo sítio no mesmo gene.

**Homólogos.** Veja **Cromossomos homólogos**; **Genes homólogos**.

**Homozigoto.** Indivíduo no qual as duas cópias de um gene são o mesmo alelo.

**Hormônio do crescimento humano (HGH).** Polipeptídeo sinalizador necessário para o crescimento normal de seres humanos; é deficiente em indivíduos com determinados tipos de nanismo.

**Hormônio.** Produto orgânico de células de uma parte do corpo que é transportado pelos líquidos corporais para outra parte onde influencia a atividade ou atua como agente coordenador.

## I

**Ilhas de CpG.** Aglomerados de citosinas e guaninas que normalmente ocorrem na região 5' dos genes humanos.

**Impressão digital de DNA.** Veja **Perfil de DNA**.

**Imprinting.** Processo que altera o estado de um gene sem alterar sua sequência de nucleotídeos; muitas vezes está associado à metilação de nucleotídeos específicos no gene. O estado alterado é estabelecido na linhagem germinativa e transmitido à prole, na qual pode persistir por toda a vida. Um gene alterado dessa maneira é denominado *imprinted*.

**Imunoglobulina.** Veja **Globulinas**.

**In situ.** Do latim, significa no lugar natural; refere-se a tratamentos experimentais realizados em células ou tecidos, e não em extratos deles.

**In vitro.** Do latim, significa “dentro do vidro”; processos biológicos induzidos experimentalmente fora do organismo em um tubo de ensaio ou outro recipiente.

**In vivo.** Do latim, significa “dentro do organismo vivo”.

**Incompatibilidade gametofítica.** Fenômeno botânico controlado pelo locus S complexo no qual um grão de pólen não é capaz de fertilizar um óvulo produzido por um vegetal que tem alelo S igual ao do grão de pólen. Por exemplo, o pólen  $S_1$  não pode fertilizar um óvulo produzido por uma planta  $S_1/S_2$ .

**Indução.** Processo de ativar a expressão de um gene ou conjunto de genes por um indutor.

**Indutor.** Substância de baixo peso molecular que se liga a um repressor, produzindo um complexo que não é mais capaz



de se ligar ao operador; assim, a presença do indutor ativa a expressão do(s) gene(s) controlado(s) pelo operador.

**Inibição do produto final.** Veja **Inibição por feedback**.

**Inibição por feedback (ou inibição pelo produto final).** O produto final acumulado de uma via bioquímica interrompe a síntese desse produto. Um metabólito final de uma via de síntese regula a síntese em uma etapa anterior dessa mesma via.

**Inibidor.** Toda substância ou objeto que retarda uma reação química; um gene de efeito maior ou modificador que interfere em uma reação.

**Iniciação (da síntese de DNA, RNA ou proteínas).** Incorporação da primeira subunidade (nucleotídeo ou aminoácido) durante a síntese de uma macromolécula (DNA, RNA ou polipeptídeo).

**Iniciador (primer).** Sequência curta de nucleotídeos que tem um 3'OH reativo e pode iniciar a síntese de DNA ao longo de um molde.

**Iniciador (primer) de RNA.** Segmento de RNA curto (10 a 60 nucleotídeos) usado para iniciar a síntese de um novo filamento de DNA; sintetizado pela enzima DNA primase.

**Inteína.** Sequência curta de aminoácidos em um produto primário da tradução que é capaz de excisar a si própria do polipeptídeo.

**Interação.** Em estatística, efeito que não pode ser explicado pela soma da ação dos fatores de contribuição; um desvio da adição estrita.

**Interações hidrofóbicas.** Associação de grupos apolares entre si quando em soluções aquosas em razão da sua insolubilidade em água.

**Intercruzamento.** Cruzamento entre híbridos  $F_1$  derivados de um cruzamento entre pais de duas linhagens.

**Intérfase.** Estágio do ciclo celular em que a célula não está se dividindo; estágio metabólico durante o qual há replicação do DNA; o estágio que começa depois da telófase de uma divisão e se estende até o início da prófase da próxima divisão.

**Interferência por RNA (RNAi).** Fenômeno no qual o RNA bifilamentar impede a expressão de um gene que é homólogo, ao menos em parte, ao RNA.

**Interferência.** *Crossing over* em um ponto que reduz a chance de outro *crossing over* próximo; detectada por estudo do padrão de *crossing over* com três ou mais genes ligados.

**Intersexo.** Organismo que apresenta características sexuais secundárias intermediárias entre os sexos masculino e feminino; tipo que tem algumas características fenotípicas dos sexos masculino e feminino.

**Íntrons.** Sequências intercalares de bases de DNA nos genes eucarióticos que não são representadas no transcrito de RNA maduro porque são cortadas do transcrito de RNA primário.

**Invariável.** Constante, inalterável, geralmente se refere à porção de uma molécula que é igual em diferentes espécies.

**Inversão.** Rearranjo que inverte a ordem de um arranjo linear de genes em um cromossomo.

**Inversão paracêntrica.** Inversão que está totalmente dentro de um braço de um cromossomo e não inclui o centrômero.

**Inversão pericêntrica.** Inversão que inclui o centrômero, portanto abrange os dois braços de um cromossomo.

**Isoalelos.** Diferentes formas de um gene que produzem fenótipo igual ou fenótipos muito semelhantes.

**Isocromossomo.** Um cromossomo com dois braços idênticos e genes idênticos. Os braços são imagens espelhadas um do outro.

**Isoforma.** Membro de uma família de proteínas intimamente relacionadas – proteínas que têm algumas sequências de aminoácidos em comum e outras diferentes.

**Isótopo radioativo.** Isótopo (forma de um átomo) instável que emite radiação ionizante.

## L

**Lamela.** Estrutura, placa ou vesícula formada por duas membranas paralelas.

**Leptóteno (leptonema).** Estágio da meiose imediatamente antes da sinapse no qual os cromossomos são observados como estruturas filiformes, finas e únicas (mas são, na verdade, duplas porque já houve replicação do DNA).

**Letal equilibrado.** Mutações letais em diferentes genes do mesmo par de cromossomos que são mantidas em repulsão por causa da ligação próxima ou supressão do *crossing over*. Em uma população fechada, sobrevivem apenas os heterozigotos *trans* ( $l_1 + l_2$ ) para as mutações letais.

**Ligação.** A união de duas ou mais moléculas de DNA por ligações covalentes.

**Ligação covalente.** Ligação na qual um par de elétrons é igualmente compartilhado por prótons de dois átomos adjacentes.

**Ligação gênica.** Tendência de diferentes genes de serem herdados juntos porque estão localizados no mesmo cromossomo.

**Ligação peptídica.** Ligação química que une as subunidades de aminoácidos nas proteínas.

**Ligação sexual.** Associação ou ligação de uma característica hereditária ao sexo; o gene está em um cromossomo sexual, geralmente o X; muitas vezes é empregado como sinônimo de ligação ao X.

**Ligações de hidrogênio.** Interações fracas entre átomos eletro-negativos e átomos de hidrogênio (eletropositivos) que estão ligados a outros átomos eletronegativos.

**Ligações iônicas.** Atrações entre grupos químicos de cargas opostas.

**Ligador (DNA).** Dupla hélice de DNA desprotegida que conecta nucleossomos adjacentes.

**Ligante.** Molécula que pode se ligar a outra molécula no interior das células ou sobre elas.

**Ligase.** Enzima que une as extremidades de dois filamentos de ácido nucleico

**Limitado ao sexo.** Expressão de uma característica em apenas um sexo. Exemplos: produção de leite por mamíferos; chifres em carneiros Rambouillet; produção de ovos por galinhas.

**LINE (elementos nucleares intercalados longos).** Famílias de elementos transponíveis moderadamente repetitivos longos (comprimento médio = 6.500 pb) em eucariotos.

**Linfócito.** Uma classe geral de leucócitos que são componentes importantes do sistema imune de animais vertebrados.

**Linfócitos B (células B).** Importante classe de células que amadurecem na medula óssea e são as principais responsáveis



pela resposta imune mediada por anticorpos ou humoral; dão origem aos plasmócitos produtores de anticorpos e a algumas outras células do sistema imune.

**Linfócitos T (células T).** Células que se diferenciam no timo e são as principais responsáveis pela resposta imune celular ou mediada por células T.

**Linhagem endogâmica.** Linhagem produzida por muitas gerações de endogamia sistemática, por exemplo, por autofertilização repetida ou por cruzamento repetido de irmãos completos.

**Linhagem germinativa.** O tecido que produz os gametas.

**Lípido.** Molécula composta de ácidos graxos e triglicerídios.

**Líquido amniótico.** Líquido contido na bolsa amniótica de vertebrados superiores e que contém células do embrião (não da mãe). Tanto o líquido quanto as células são usados para diagnóstico de anormalidades genéticas do embrião ou do feto.

**Lise.** Ruptura de uma célula pela destruição da membrana celular após infecção viral.

**Lisossomo.** Pequena organela celular delimitada por membrana que contém enzimas destinadas à degradação de macromoléculas.

**Loci de característica quantitativa (QTL).** Dois ou mais genes que afetam uma única característica quantitativa.

**Locus (pl. loci).** Posição fixa em um cromossomo ocupada por determinado gene ou um de seus alelos.

## M

**Macromolécula.** Molécula grande; termo usado para identificar moléculas de proteínas e ácidos nucleicos.

**Mapa citológico.** Diagrama de um cromossomo baseado na coloração diferencial – o “padrão de bandas” – ao longo de sua extensão.

**Mapa de ligação gênica.** Diagrama linear ou circular que mostra as posições relativas de genes em um cromossomo obtido por análise genética.

**Mapa de restrição.** Mapa físico linear ou circular de uma molécula de DNA que mostra os sítios clivados por diferentes enzimas de restrição.

**Mapa físico.** Diagrama de um cromossomo ou molécula de DNA com distâncias apresentadas em pares de bases, quilobases ou megabases.

**Mapa genético.** Diagrama de um cromossomo com distâncias baseadas em frequências de recombinação – centimorgans.

**Mapeamento de híbridos por radiação.** Uso de células híbridas de seres humanos e roedores que contêm fragmentos de cromossomos humanos (produzidos por irradiação) fundidos a cromossomos de roedores para determinar as relações de ligação de genes humanos.

**Marcação por transpósom.** Inserção de um elemento transponível dentro ou perto de um gene, marcando esse gene com uma sequência de DNA conhecida.

**Matriz aberta de leitura (ORF).** Segmento de DNA contendo as sequências necessárias para codificar um polipeptídeo. O transcrito de RNA de uma ORF começa com um códon de iniciação da tradução, seguido por uma sequência de códons especificadores de aminoácidos, e termina com um códon de término da tradução. Presume-se que a ORF codifique um polipeptídeo.

**Matriz de leitura.** Série de trinucleotídeos posicionados em sequência no sítio A do ribossomo durante a tradução de um mRNA; ainda, a sequência de trincas de pares de nucleotídeos no DNA que correspondem a esses códons no mRNA.

**Mecanismo de controle negativo.** Mecanismo em que há necessidade de proteína(s) reguladora(s) para desativar a expressão gênica.

**Mecanismo de controle positivo.** Mecanismo em que há necessidade de proteína(s) reguladora(s) para ativar a expressão gênica.

**Média.** Média aritmética; a soma de todas as medidas ou valores em uma amostra dividida pelo tamanho da amostra.

**Mediana.** Em um conjunto de medidas, o valor central acima e abaixo do qual há um número igual de medidas.

**Megaspório.** Célula grande isolada produzida no fim da meiose nos tecidos reprodutivos femininos de vegetais.

**Meiose.** Processo pelo qual o número de cromossomos de uma célula reprodutiva é reduzido à metade do número diploide ( $2n$ ) ou somático; resulta na formação de gametas em animais ou de esporos em vegetais; importante criadora de variabilidade por meio de recombinação.

**Melanina.** Pigmento castanho ou preto.

**Membrana.** Estrutura macromolecular composta de lipídios e proteínas que circunda uma célula ou algumas organelas dentro de uma célula, como as mitocôndrias e os cloroplastos; também é um componente do retículo endoplasmático celular.

**Mesoderma.** Camada germinativa média que se forma no embrião inicial de animais e dá origem a partes como osso e tecido conjuntivo.

**Metabolismo.** Soma de todos os processos químicos em células vivas em que há produção e uso de energia.

**Metáfase.** Estágio da divisão celular no qual os cromossomos são mais bem definidos e dispostos em uma placa equatorial; estágio depois da prófase e antes da anáfase.

**Metáfase I.** Estágio da primeira divisão meiótica em que os cromossomos homólogos duplicados emparelhados se condensam e se reúnem no plano equatorial da célula.

**Metáfase II.** Estágio da segunda divisão meiótica em que os cromossomos duplicados se reúnem no plano equatorial da célula.

**Metafêmea (superfêmea).** Em *Drosophila*, fêmea anormal, geralmente estéril, com excesso de cromossomos X em comparação com os conjuntos de autossomos (p. ex., XXX; AA).

**Metástase.** Disseminação de células cancerosas para órgãos antes não afetados.

**Metilação (de DNA e RNA).** Acréscimo de um ou mais grupos metil ( $-CH_3$ ) a um ou mais dos nucleotídeos em um ácido nucleico.

**Método CIB.** Uso de um cromossomo X especial de *Drosophila*, que tem uma mutação causadora de olhos em forma de barra e uma mutação letal recessiva dentro de uma grande inversão, para detectar novas mutações letais recessivas ligadas ao X. H. J. Muller usou esse cromossomo para demonstrar que os raios X são mutagênicos. Veja também **Cromossomo CIB**.

**Micélio.** Filamentos que constituem a parte vegetativa do talo dos fungos.

**Microarranjo.** Membrana ou outro suporte sólido contendo milhares de oligonucleotídeos ou sondas de hibridização



- de ácido nucleico para uso na detecção de DNA ou RNA complementares.
- Micro-RNA.** Veja **RNA de interferência curto**.
- Microsporo.** Um dos quatro produtos finais da meiose nos tecidos reprodutivos masculinos dos vegetais.
- Microsatélite.** Veja **Repetição curta em série (STR)**.
- Microtúbulos.** Filamentos ocos no citoplasma constituem uma parte do aparelho locomotor de uma célula móvel; componente do fuso mitótico.
- Mitocôndrias.** Organelas do citoplasma de células vegetais e animais onde ocorre a fosforilação oxidativa para produção de ATP.
- Mitose.** Disjunção de cromossomos duplicados e divisão do citoplasma para produzir duas células-filhas geneticamente idênticas.
- Modelo óperon.** Mecanismo de controle negativo proposto por Jacob e Monod em 1961 para explicar a regulação coordenada de conjuntos cotranscritos de genes estruturais. O mecanismo inclui um gene regulador codificador de um repressor que controla a transcrição do conjunto de genes mediante ligação a uma região operadora e bloqueio da transcrição pela RNA polimerase.
- Modelo.** Descrição matemática de um fenômeno biológico.
- Modificação tautomérica.** Transferência de um átomo de hidrogênio de uma posição para outra em uma molécula orgânica.
- Modificador (gene modificador).** Gene que afeta a expressão de algum outro gene.
- Molde.** Padrão ou modelo. O DNA armazena informações codificadas e atua como um modelo ou molde a partir do qual as informações são copiadas em filamentos complementares de DNA ou transcritos para o RNA mensageiro.
- Molécula de DNA recombinante.** Molécula de DNA construída *in vitro* por união total ou parcial de duas moléculas diferentes de DNA.
- Molécula efetora.** Molécula que influencia o comportamento de uma molécula reguladora, como uma proteína repressora, assim influenciando a expressão gênica.
- Mono-híbrido.** Prole de dois homozigotos que diferem um do outro pelos alelos presentes em apenas um *locus* gênico.
- Monômero.** Molécula única que pode se combinar a outras para formar estruturas mais complexas.
- Monoploide.** Organismo ou célula que tem apenas um conjunto de cromossomos ou um genoma (número *n* de cromossomos).
- Monossômico.** Célula ou organismo diploide em que falta um cromossomo de seu próprio complemento (fórmula cromossômica  $2n - 1$ ). Um caso específico desse distúrbio é denominado monossomia.
- Morfógeno.** Substância que estimula o desenvolvimento da forma ou estrutura de um organismo.
- Morfologia.** Estudo da forma de um organismo; história do desenvolvimento de estruturas visíveis e a relação comparativa de estruturas semelhantes em diferentes organismos.
- Mosaico.** Organismo ou parte de um organismo composto de células de diferentes genótipos.
- Mosaico sexual.** Veja **Ginandromorfo**.
- Motilidade.** Movimento celular, geralmente pela ação de estruturas especializadas, como cílios e flagelos.
- mtDNA.** DNA das mitocôndrias.
- Mutação.** Alteração do DNA em um *locus* específico de um organismo. O termo é usado de modo abrangente e inclui mutações pontuais com alteração envolvendo um único gene, bem como uma alteração cromossômica.
- Mutação com ganho de função.** Mutação que dota um produto gênico de uma nova função.
- Mutação com perda de função.** Mutação que compromete ou suprime a expressão gênica ou a função de um produto gênico.
- Mutação de efeito materno.** Mutação que causa fenótipo mutante na prole de um indivíduo do sexo feminino portador da mutação; no entanto, o próprio portador do sexo feminino pode não apresentar o fenótipo mutante.
- Mutação dominante negativa.** Alelo mutante de um gene que interfere na função de um alelo selvagem de maneira que indivíduos heterozigotos para os alelos mutante e selvagem tenham um fenótipo mutante.
- Mutação espontânea.** Mutação que ocorre sem causa conhecida (cf. **Mutação induzida**).
- Mutação germinativa.** Mutação que ocorre nas células reprodutivas (células germinativas) do corpo e é transmitida à prole (cf. **Mutação somática**).
- Mutação homeótica.** Mutação que causa o desenvolvimento de uma parte do corpo em posição imprópria em um organismo; por exemplo, uma mutação em *Drosophila* que causa o desenvolvimento de pernas na cabeça em lugar das antenas.
- Mutação induzida.** Mutação resultante da exposição de um organismo a um agente químico ou físico que causa alterações na estrutura do DNA ou RNA (cf. **Mutação espontânea**).
- Mutação knockout.** Mutação que suprime totalmente a função de um gene.
- Mutação letal condicional.** Mutação que é letal em uma série de condições ambientais – condições restritivas –, mas é viável em outra série de condições ambientais – condições permissivas.
- Mutação letal recessiva.** Forma mutante de um gene que provoca a morte de um organismo homozigoto para ele.
- Mutação neutra.** Mutação que altera a sequência de nucleotídeos de um gene, mas não afeta a aptidão do organismo.
- Mutação nula.** Mutação que suprime a expressão de um gene. (Veja também **Amórfico**.)
- Mutação por inserção.** Mutação causada pela inserção de DNA estranho como um elemento transponível ou o T-DNA do plasmídeo Ti de *Agrobacterium tumefaciens*.
- Mutação por mudança da matriz de leitura.** Mutação que modifica a matriz de leitura de um mRNA, seja por inserção ou por deleção de nucleotídeos.
- Mutação sem sentido.** Mutação que troca um códon especificador de um aminoácido por um códon de término.
- Mutação somática.** Mutação que ocorre nas células não reprodutivas (células somáticas) do corpo e não é transmitida à prole (cf. **Mutação germinativa**).
- Mutação supressora.** Mutação que anula parcial ou totalmente o efeito fenotípico de outra mutação.
- Mutações de sentido trocado.** Mutação que troca um códon que especifica um aminoácido por um códon que especifica outro aminoácido.



**Mutações pontuais.** Alterações ocorridas em sítios específicos nos genes. Incluem substituições de pares de nucleotídeos e inserção ou deleção de um ou alguns pares de nucleotídeos.

**Mutágeno.** Agente ambiental, físico ou químico, capaz de induzir mutações.

**Mutante.** Célula ou organismo individual que apresenta uma alteração produzida por mutação; um gene modificado.

**Mutante sensível ao supressor.** Organismo capaz de crescer na presença de um segundo fator genético – um supressor –, mas não na ausência desse fator.

**Mutante termossensível.** Organismo capaz de crescer em determinada temperatura, mas não em outra.

## N

**Não autônomo.** Termo designativo das unidades biológicas que não têm ação independente; essas unidades necessitam da assistência de outra unidade ou “auxiliar” (cf. **Autônomo**).

**Não disjunção.** Ausência de disjunção ou separação de cromossomos homólogos em mitose ou meiose, cuja consequência é o excesso de cromossomos em algumas células-filhas e a escassez em outras. Exemplos: na meiose, os dois cromossomos de um par vão para o mesmo polo e o outro polo não recebe nenhum deles; na mitose, as duas cromátides-irmãs vão para o mesmo polo.

**Northern blot.** Transferência de moléculas de RNA de gel de eletroforese para membrana de celulose ou náilon por ação capilar.

**Nuclease.** Enzima que catalisa a degradação de ácidos nucleicos.

**Núcleo.** Parte de uma célula eucariótica que contém os cromossomos; separada do citoplasma por uma membrana.

**Núcleo.** Bolsa esférica, rica em RNA, existente no núcleo de células metabólicas; está associado ao organizador nucleolar; local de armazenamento de ribossomos e de precursores de ribossomos.

**Nucleoproteína.** Proteína conjugada constituída de ácido nucleico e proteína; material de que são feitos os cromossomos.

**Nucleosídeo.** Substância orgânica constituída de uma base unida por ligação covalente à ribose ou desoxirribose.

**Nucleossomo, centro do nucleossomo.** A subunidade de cromatina resistente à nuclease constituída de aproximadamente 146 nucleotídeos de DNA enrolado em 1,65 volta de super-hélice negativa ao redor de um octâmero de histonas – duas moléculas de cada uma das histonas H2a, H2b, H3 e H4.

**Nucleotídeo.** Subunidade de moléculas de DNA e RNA contendo um grupo fosfato, um açúcar e uma base nitrogenada orgânica.

**Nullissômico.** Célula ou organismo diploide que não tem os dois membros de um par de cromossomos (fórmula cromossômica  $2n - 2$ ).

## O

**Octaploide.** Célula ou organismo com oito genomas ou conjuntos de cromossomos (número de cromossomos  $8n$ ).

**Oncogene.** Gene capaz de causar transformação cancerosa em células animais em cultura e formação de tumor nos próprios animais; gene que promove a divisão celular.

**Operador.** Parte de um operon que controla a expressão de um ou mais genes estruturais, servindo como sítio de ligação para uma ou mais proteínas reguladoras.

**Operon.** Grupo de genes que constituem uma unidade reguladora ou de controle. A unidade inclui um operador, um promotor e genes estruturais.

**Ordem (no código genético).** Há dois tipos de ordem no código genético: (1) múltiplos códons para determinado aminoácido geralmente diferem apenas na terceira posição, e (2) os códons para aminoácidos com propriedades químicas semelhantes estão intimamente relacionados.

**Ordenada.** Eixo vertical de um gráfico.

**Organela.** Parte especializada de uma célula que tem uma ou mais funções específicas (p. ex., o cílio de um protozoário).

**Organismos-modelo.** Vegetais, animais e micróbios usados rotineiramente em análise genética.

**Organização Genoma Humano (HUGO).** Grupo internacional de cientistas criado para coordenar o sequenciamento e o mapeamento do genoma humano.

**Organizador.** Indutor; substância química em um sistema vivo que determina o destino no desenvolvimento de determinadas células ou grupo de células.

**Organizador nucleolar (ON).** Segmento cromossômico contendo genes que controlam a síntese de RNA ribossômico, localizado na constrição secundária de alguns cromossomos.

**Origem da replicação.** Sítio ou sequência de nucleotídeos em um cromossomo ou uma molécula de DNA onde se inicia a replicação.

**Ortólogos.** Veja **Genes ortólogos**.

**Ovário.** Parte alargada do pistilo de uma flor que contém os óvulos; o órgão reprodutivo ou a gônada feminina em animais.

**Ovócito.** Célula germinativa feminina; a célula que passa por duas divisões meióticas (ovogênese) para dar origem ao óvulo. Ovócito primário – antes da conclusão da primeira divisão meiótica; ovócito secundário – depois da conclusão da primeira divisão meiótica.

**Ovócito (oosfera).** Célula germinativa produzida por um organismo feminino.

**Ovócito secundário.** Veja **Ovócito**.

**Ovogênese.** Formação do ovócito em animais.

**Ovogônia.** Célula germinativa do animal de sexo feminino antes do início da meiose.

**Óvulo (vegetal).** O macrosporângio de uma angiosperma que se transforma em semente. Abrange o nucelo e o tegumento.

## P

**P.** Símbolo que designa a geração parental ou os pais de determinado indivíduo.

**Palíndromo.** Segmento de DNA no qual a leitura da sequência de pares de bases é igual nos dois sentidos a partir de um ponto central de simetria.

**Panmixia.** Cruzamento aleatório em uma população.

**Paquíteno (paquinema).** Estágio do meio da prófase meiótica imediatamente depois do zigóteno e antes do diplóteno. Em preparações microscópicas favoráveis, os cromossomos são visíveis como pares de filamentos longos. Raramente, é possível detectar quatro cromátides.



**Parálogos.** Veja **Genes parálogos**.

**Parâmetro.** Valor ou constante com base em toda uma população (cf. **Estatística**).

**Parental.** Relativo às linhagens fundadoras usadas em um cruzamento; que tem as características dessas linhagens. Em uma série de cruzamentos, a geração parental é simbolizada pela letra P.

**Partenogênese.** Desenvolvimento de um novo indivíduo a partir de um gameta feminino não fertilizado.

**Paterno.** Relativo ao pai.

**Patógeno.** Organismo causador de doença.

**PCR.** Veja **Reação em cadeia da polimerase**.

**Penetrância.** Porcentagem de indivíduos que apresentam determinado fenótipo entre os capazes de apresentá-lo.

**Peptidil transferase.** Atividade enzimática – presente na subunidade grande do ribossomo – que catalisa a formação de ligações peptídicas entre aminoácidos durante a tradução.

**Peptídio.** Substância contendo aminoácidos; unidade de decomposição ou de formação no metabolismo de proteínas.

**Pequenas ribonucleoproteínas nucleares (snRNP).** Complexos de RNA-proteína que são componentes dos espliceossomos.

**Pequeno RNA nuclear (snRNA).** Pequenas moléculas de RNA localizadas nos núcleos de células eucarióticas; a maioria dos snRNAs compõe os espliceossomos que excisam íntrons dos pré-mRNA.

**Perfil de DNA (impressão digital de DNA).** Padrão registrado de polimorfismos de DNA.

**Peroxisomo.** Organela subcelular que contém enzimas que participam da degradação de ácidos graxos e aminoácidos.

**Pintura cromossômica.** Estudo da organização e evolução de cromossomos por hibridização *in situ* com uso de sondas de DNA marcadas com corantes fluorescentes que emitem luz em diferentes comprimentos de onda.

**Pirimidina.** Base nitrogenada de anel simples presente em ácidos nucleicos; a citosina e a timina geralmente estão presentes no DNA, ao passo que a uracila costuma substituir a timina no RNA.

**Pistilo.** Órgão central das flores que contém o ovário.

**Placa equatorial.** Figura formada pelos cromossomos no centro (plano equatorial) do fuso na mitose.

**Placa metafásica.** Plano equatorial em que se reúnem os cromossomos duplicados de uma célula durante a metáfase da mitose.

**Plaqueamento em réplica.** Procedimento de duplicação de colônias bacterianas cultivadas em meio de ágar em uma placa de Petri para meio de ágar em outra placa de Petri.

**Plasmídio.** Determinante hereditário extracromossômico que existe em estado autônomo e é transferido de maneira independente dos cromossomos.

**Plasmídio R conjugativo.** Molécula de DNA circular que pode ser transferida de uma bactéria para outra durante a conjugação.

**Plasmídio Ti.** O grande plasmídio em *Agrobacterium tumefaciens*. É responsável pela indução de tumores em vegetais com a doença galha-da-coroa e é um importante vetor para transferência de genes para os vegetais, principalmente as dicotiledôneas.

**Plasmócitos.** Leucócitos produtores de anticorpos derivados de linfócitos B.

**Plastídio.** Corpo citoplasmático encontrado nas células de vegetais e de alguns protozoários. Os cloroplastos, por exemplo, produzem clorofila, que participa da fotossíntese.

**Pleiotropia.** Condição na qual um só gene influencia mais de uma característica.

**Pluripotente.** Adjetivo que designa células com potencial de se diferenciar em muitos tipos diferentes.

**Poliadenilação.** Acréscimo de caudas poli(A) a transcritos de genes eucarióticos (RNA).

**Polidactilia.** Ocorrência de número maior que o habitual de dedos das mãos ou dos pés.

**Poligene.** Um de muitos genes implicados na herança quantitativa.

**Poliligador (sítio de clonagem múltipla).** Segmento de DNA que contém um conjunto de sítios de clivagem de enzima de restrição única.

**Polimerase poli(A).** Enzima que acrescenta as caudas poli(A) às terminações 3' dos transcritos (RNA) de genes eucarióticos.

**Polimerase.** Enzima que catalisa a síntese de DNA ou RNA.

**Polimerização.** União química de duas ou mais moléculas do mesmo tipo para formar uma nova substância que tem elementos iguais em proporções iguais, porém maior peso molecular e diferentes propriedades físicas.

**Polímero.** Substância constituída de muitas subunidades menores; decorrente do processo de polimerização.

**Polimorfismo de nucleotídeo único (SNP).** Um par de bases no DNA que varia em uma população.

**Polimorfismo do comprimento do fragmento de restrição (RFLP).** Existência de duas ou mais variantes genéticas detectáveis por observação de fragmentos de DNA genômico obtidos por digestão do DNA com uma enzima de restrição. Em geral, os fragmentos de DNA são fracionados por eletroforese, transferidos para uma membrana por *Southern blotting* e, depois, visualizados por autorradiografia após hibridização com sonda de DNA marcada.

**Polimorfismo equilibrado.** Dois ou mais tipos de indivíduos mantidos na mesma população reprodutiva por um mecanismo de seleção.

**Polimorfismo silencioso.** Variante de DNA que não altera a sequência de aminoácidos de uma proteína.

**Polimorfismo.** Existência de duas ou mais variantes em uma população de indivíduos em que no mínimo duas variantes têm frequências acima de 1%.

**Polinucleotídeo.** Sequência linear de nucleotídeos unidos no DNA ou RNA.

**Poli-peptídio.** Molécula linear com dois ou mais aminoácidos e um ou mais grupos peptídicos. São denominados dipeptídios, tripeptídios, e assim por diante, de acordo com o número de aminoácidos presentes.

**Poli-ploide.** Organismo com mais de dois conjuntos de cromossomos (diploide  $2n$ ) ou genomas – por exemplo, triploide ( $3n$ ), tetraploide ( $4n$ ), pentaploide ( $5n$ ), hexaploide ( $6n$ ), heptaploide ( $7n$ ), octaploide ( $8n$ ).

**Ponto de verificação.** Mecanismo que impede o avanço durante o ciclo celular eucariótico.

**Pool de genes.** A soma de todos os alelos diferentes nos membros reprodutores de uma população em dado momento.



- População.** Grupo de organismos de um tipo; grupo de inter-cruzamento de vegetais ou animais; o amplo grupo do qual se pode retirar uma amostra.
- População (efetiva).** Membros reprodutores da população.
- População mendeliana.** Unidade natural de inter-cruzamento de vegetais com reprodução sexuada ou animais que têm um *pool* de genes em comum.
- População pan-mítica.** População em que o cruzamento é aleatório.
- Portador.** Indivíduo que tem um alelo recessivo não expresso (ou seja, inibido por um alelo dominante).
- Pré-mRNA.** Transcrito primário de um gene eucariótico antes do processamento para produzir um mRNA.
- Pressão de mutação.** Taxa de mutação constante que acrescenta genes mutantes a uma população; ocorrências repetidas de mutações em uma população.
- Pressão de seleção.** Eficácia das diferenças de sobrevivência e reprodução na modificação da frequência de alelos em uma população.
- Primossomo.** Complexo de replicação proteico que catalisa a iniciação dos fragmentos de Okazaki durante a síntese descontínua do DNA. Contém atividades de DNA primase e DNA helicase.
- Princípio de Hardy-Weinberg.** Relação matemática que permite prever as frequências de genótipos em uma população a partir das frequências dos alelos constituintes; uma consequência do cruzamento aleatório.
- Princípio do fundador.** Possibilidade de divergência genética de uma população nova, pequena e isolada porque os indivíduos fundadores são uma amostra aleatória de uma grande população principal.
- Probabilidade.** Frequência de ocorrência de um evento.
- Probabilidade binomial.** Frequência associada à ocorrência de um resultado em um experimento que tem apenas dois resultados possíveis, como cara ou coroa ao jogar uma moeda.
- Probando.** Indivíduo de uma família no qual se identifica pela primeira vez uma característica hereditária.
- Procarioto.** Membro de um grande grupo de organismos (entre eles bactérias e cianobactérias) que não têm núcleo celular verdadeiro e não sofrem mitose.
- Prófago (provírus).** Genoma de um bacteriófago temperado integrado ao cromossomo de uma bactéria lisogênica e replicado com o cromossomo do hospedeiro.
- Prófase.** Estágio da mitose entre a intérfase e a metáfase. Durante essa fase, o centríolo se divide e os dois centríolos resultantes se afastam. Há espiralamento de cada filamento de DNA produzido por replicação na intérfase, e o cromossomo é duplo no sentido longitudinal, exceto na região do centrômero. Cada cromossomo parcialmente separado é denominado cromátide. As duas cromátides de um cromossomo são cromátides-irmãs.
- Prófase I.** Estágio da primeira divisão meiótica em que há condensação dos cromossomos duplicados e emparelhamento de cromossomos homólogos.
- Prófase II.** Estágio da segunda divisão meiótica em que os cromossomos duplicados se condensam e se preparam para se deslocar até o plano equatorial da célula.
- Progérias.** Doenças hereditárias caracterizadas por envelhecimento prematuro.
- Projeto Genoma Humano.** Enorme esforço internacional para mapear e sequenciar todo o genoma humano.
- Promotor.** Sequência nucleotídica à qual a RNA polimerase se liga para iniciar a transcrição; também é uma substância química que estimula a transformação de células benígnas em células cancerosas.
- Protaminas.** Pequenas proteínas básicas que substituem as histonas nos cromossomos de alguns gametas masculinos.
- Protease.** Toda enzima que hidrolisa proteínas.
- Proteína.** Macromolécula constituída de um a vários polipeptídios. Cada polipeptídio é constituído de uma cadeia de aminoácidos unidos por ligações peptídicas.
- Proteína ativadora do catabolismo (CAP).** Proteína reguladora positiva que, na presença de AMP cíclico (cAMP), liga-se às regiões promotoras de óperons e estimula sua transcrição. A CAP e o cAMP, quando presentes, garantem o uso de glicose como fonte de carbono em vez de fontes de energia menos eficientes como a lactose, a arabinose e outros açúcares. Quando presente, a glicose impede a síntese de cAMP e, portanto, a ativação da transcrição por CAP/cAMP.
- Proteína de fusão.** Polipeptídio constituído de um gene recombinante que contém porções de dois ou mais genes diferentes. Os diferentes genes são unidos de modo que suas sequências codificadoras estejam na mesma matriz de leitura.
- Proteína de ligação ao DNA unifilamentar.** Proteína que reveste filamentos únicos de DNA, mantendo-os no estado estendido.
- Proteína fluorescente verde (GFP).** Proteína fluorescente natural sintetizada pela água-viva *Aequorea victoria*.
- Proteínas cromossômicas não histônicas.** Todas as proteínas dos cromossomos exceto as histonas.
- Proteínas de competência (Com).** Proteínas mediadoras do processo de transformação em bactérias. Sua síntese é induzida por pequenos peptídios denominados feromônios de competência.
- Proteoma.** Conjunto completo de proteínas codificadas por um genoma.
- Proteômica.** Ciência que determina as estruturas e as funções de todas as proteínas produzidas por organismos vivos.
- Proto-oncogene.** Gene celular normal que pode ser transformado em oncogene por mutação.
- Protoplasto.** Célula vegetal ou bacteriana cuja parede foi removida.
- Prototrofo.** Organismo como uma bactéria que cresce em meio mínimo.
- Provírus.** Cromossomo viral que se integrou ao genoma de um hospedeiro, procariótico ou eucariótico (cf. **Prófago**).
- Pseudogene.** Componente inativo, mas estável de um genoma, semelhante a um gene; aparentemente derivado de genes ativos por mutação.
- Purina.** Base nitrogenada de anel duplo presente em ácidos nucleicos; a adenina e a guanina são as duas purinas presentes na maioria das moléculas de DNA e RNA.

## Q

**Quiasma.** Troca visível de partes em duas cromátides de um grupo de quatro durante a primeira prófase meiótica. No



estágio diplóteno da meiose, as quatro cromátides de um bivalente estão associadas em pares, mas de tal modo que há troca de uma parte de duas cromátides. Esse ponto de "troca de partes" é o quiasma.

**Quimera (animal).** Indivíduo derivado de dois embriões por intervenção experimental.

**Quimera (vegetal).** Parte de um vegetal cuja constituição genética é diferente da constituição de outras partes do mesmo vegetal. Pode ser resultado de diferentes zigotos que crescem juntos ou de fusão artificial (enxerto); pode ser periclinal, com camadas paralelas de tecidos geneticamente diferentes, ou setorial.

**Quimiotaxia.** Atração ou repulsão de organismos por uma substância capaz de se difundir.

**Qui-quadrado.** Dado estatístico usado para testar o ajustamento dos dados às previsões de uma hipótese.

## R

**Raça.** Grupo distinguível de organismos de determinada espécie.

**Radiação ionizante.** Parte do espectro eletromagnético que resulta na produção de cargas positivas e negativas (pares de íons) em moléculas. Os raios X e os raios gama são exemplos de radiação ionizante (cf. **Radiação não ionizante**).

**Radiação não ionizante.** Parte do espectro eletromagnético que não leva à produção de cargas positivas e negativas (pares de íons) em moléculas. A luz visível e a luz ultravioleta são exemplos de radiação não ionizante (cf. **Radiação ionizante**).

**Radiação ultravioleta (UV).** A parte do espectro eletromagnético – comprimentos de onda de aproximadamente 1 a 350 nm – entre a radiação ionizante e a luz visível. A radiação UV é absorvida pelo DNA e é extremamente mutagênica para organismos unicelulares e para as células epidérmicas de organismos multicelulares.

**Rastreamento por complementação.** Pesquisa em bibliotecas de expressão de cDNA ou clones genômicos com base em sua capacidade de resgatar células hospedeiras mutantes.

**Reação autocatalítica.** Reação catalisada por um substrato sem a participação de nenhum outro agente catalítico.

**Reação em cadeia da polimerase (PCR).** Procedimento que requer vários ciclos de desnaturação, hibridização em oligonucleotídeos iniciadores e síntese de polinucleotídeos, que amplifica determinada sequência de DNA.

**Receptor.** Molécula capaz de aceitar a união de um ligante.

**Receptor de células T.** Proteína de ligação a antígeno que está localizada na superfície de células T e medeia a resposta imune celular de mamíferos. Os genes que codificam antígenos de células T são montados a partir de segmentos gênicos por processos de recombinação somática durante a diferenciação de linfócitos T.

**Recessivo.** Termo que designa um membro de um par alélico incapaz de se manifestar na presença do membro dominante.

**Recombinação.** Produção de combinações gênicas não encontradas nos pais pela distribuição de cromossomos não homólogos e *crossing over* entre cromossomos homólogos durante a meiose. Nos genes ligados, pode-se usar a frequência de recombinação para estimar a distância no mapa genético; no entanto, altas frequências (próximas de 50%) não produzem estimativas precisas.

**Recomposição (*splicing*).** Processo responsável pela ligação covalente de sequências de éxons de RNA e pela eliminação de sequências de íntrons interpostas.

**Região *vir* (de plasmídeo Ti).** Região do plasmídeo Ti de *Agrobacterium tumefaciens* que contém genes codificadores de produtos necessários para a transferência do T-DNA da bactéria para células vegetais.

**Regras de Kozak.** A sequência necessária – 5'-GCC(A ou G)CCAUGG-3' – para o início ideal da tradução no primeiro (5') AUG em mRNA eucarióticos (nome dado em homenagem a Marilyn Kozak, primeira a propor essas regras).

**Remodelagem da cromatina.** Alteração da estrutura do DNA e de suas moléculas de proteína associadas, principalmente histonas, por um complexo proteico; muitas vezes essa remodelagem implica a modificação química das histonas.

**Renaturação.** Restauração da forma nativa de uma molécula. Na bioquímica de ácidos nucleicos, esse termo geralmente se refere à formação de uma hélice bifilamentar a partir de moléculas unifilamentares complementares.

**Reparo de erros de pareamento.** Processos de reparo do DNA que corrigem pares de bases que não têm ligações de hidrogênio apropriadas.

**Reparo por excisão.** Processos de reparo do DNA em que há retirada do segmento lesado de DNA e sua substituição pela síntese de um novo filamento usando o filamento complementar de DNA como molde.

**Reparo por excisão de base.** Retirada de bases anormais ou quimicamente modificadas do DNA.

**Reparo por excisão de nucleotídeo.** Eliminação de defeitos relativamente grandes, como dímeros de timina no DNA, mediante excisão de um segmento do filamento de DNA que transpõe o defeito e a síntese de reparo por uma DNA polimerase que usa o filamento complementar como molde.

**Reparo pós-replicação.** Mecanismo dependente de recombinação para reparo do DNA lesado.

**Repetição curta em série (STR) (microssatélite).** Repetição em série altamente polimórfica de uma sequência que tem apenas dois a cinco pares de nucleotídeos de comprimento.

**Repetição em série em número variável (VNTR) (minissatélite).** Repetição em série extremamente polimórfica de uma sequência com 10 a 80 pares de nucleotídeos de comprimento.

**Repetição em série simples.** Unidade repetida em série no DNA com apenas um a seis nucleotídeos de comprimento.

**Repetição invertida.** Sequência presente duas vezes em uma molécula de DNA, mas com orientação invertida.

**Repetição terminal invertida.** Sequências de DNA idênticas ou quase idênticas em extremidades opostas de um transpósom de "cortar e colar". Uma sequência é a imagem especular invertida da outra.

**Repetições de trinucleotídeos.** Repetições consecutivas de três nucleotídeos presentes em muitos genes humanos. Em vários casos, essas repetições de trinucleotídeos sofreram expansões do número de cópias que acarretaram doenças hereditárias.

**Repetições terminais longas.** Sequências de DNA idênticas ou quase idênticas em extremidades opostas de um retrovírus ou de um elemento semelhante a um retrovírus integrado. Em geral, essas sequências têm comprimento mínimo de 300 pares de bases. Sigla: LTR (*long terminal repeats*).



- Replicação.** Processo de duplicação realizado por cópia de um molde (p. ex., reprodução no nível de DNA).
- Replicação contínua.** Síntese de um novo filamento de DNA pelo acréscimo sequencial de nucleotídeos à terminação 3'-OH do filamento. Característica da síntese do filamento contínuo (líder) – o filamento que cresce na direção geral 5' → 3'.
- Replicação descontínua.** A síntese de um novo filamento de DNA pela formação de segmentos curtos de DNA (fragmentos de Okazaki) que são, em seguida, unidos pela DNA ligase. Característica da síntese do filamento descontínuo (atrasado) – o filamento em crescimento na direção geral 3' → 5'.
- Replicação por círculo rolante.** Mecanismo de replicação de moléculas circulares de DNA nas quais um filamento parental de DNA é clivado na origem de replicação enquanto o outro filamento permanece intacto. A terminação 5' do filamento clivado é desenrolada e replicada de modo descontínuo, enquanto a replicação contínua do outro filamento ocorre na terminação 3', usando o filamento circular intacto como molde.
- Replicação semiconservativa.** Replicação de DNA por mecanismo no qual os filamentos parentais são conservados (permanecem intactos) e servem como moldes para a síntese de novos filamentos complementares.
- Réplicon.** Unidade de replicação. Em bactérias, os réplicons estão associados a segmentos da membrana celular que controlam a replicação e a coordenam com a divisão celular.
- Replissomo.** O aparelho de replicação completo – presente em uma forquilha de replicação – que realiza a replicação semiconservativa do DNA.
- Repressão.** Processo de desativar a expressão de um gene ou conjunto de genes em resposta a algum sinal.
- Repressão catabólica.** Redução mediada por glicose das taxas de transcrição de óperons que especificam enzimas participantes das vias catabólicas (como o óperon *lac*).
- Repressão coordenada.** Regulação correlacionada dos genes estruturais de um óperon por uma molécula que interage com a sequência do operador.
- Repressor.** Proteína que se liga ao DNA e desativa a expressão gênica.
- Reprodução assexuada.** Todo processo de reprodução em que não há formação e união de gametas dos diferentes sexos ou cruzamentos.
- Reprodução sexuada.** Reprodução que abrange a formação de células germinativas maduras (i. e., ovócitos/oosferas e espermatozoides/anterozoides).
- Repulsão (configuração *trans*).** Condição na qual um heterozigoto duplo recebeu um alelo mutante e um alelo selvagem de cada um dos pais; por exemplo,  $a+/a + \times +b/+b$  produz  $a+/+b$  (cf. **Acoplamento**).
- Resposta à seleção.** Na reprodução de vegetais e animais, a diferença entre a média dos indivíduos selecionados para serem pais e a média da sua prole.
- Resposta SOS.** Síntese de um conjunto completo de proteínas de reparo, recombinação e replicação do DNA em bactérias que sofrem graves danos do DNA (p. ex., após exposição à luz UV).
- Retículo endoplasmático.** Rede de membranas no citoplasma à qual os ribossomos se aderem.
- Reticulócito.** Hemácia jovem.
- Retrocruzamento.** Cruzamento entre um híbrido  $F_1$  e um de seus pais. A prole desse cruzamento é denominada geração de retrocruzamento. (Veja também **Cruzamento-teste**.)
- Retroelemento.** Qualquer um dos retrovírus integrados ou dos elementos transponíveis semelhantes a eles.
- Retromutação.** Segunda mutação, ocorrida no mesmo sítio gênico que a primeira, que restaura a sequência de nucleotídeos do tipo selvagem.
- Retroposon.** Elemento transponível que cria novas cópias por transcrição reversa de RNA em DNA, mas não tem as sequências de repetições terminais longas.
- Retrotransposon.** Elemento transponível que cria novas cópias por transcrição reversa de RNA em DNA.
- Retrovírus.** Vírus que armazena suas informações genéticas em RNA e as replica usando a transcriptase reversa para sintetizar uma cópia em DNA do genoma de RNA.
- Reversão (mutação reversa).** Restituição de um gene mutante à condição selvagem ou, no mínimo, a uma forma que produz o fenótipo selvagem; de modo geral, o aparecimento de uma característica expressa por ancestral remoto.
- Revisão (*proofreading*).** Varredura do DNA por enzimas à procura de defeitos estruturais, como erro de pareamento de bases.
- RFLP.** Veja **Polimorfismo do comprimento do fragmento de restrição**.
- Ribonuclease (RNase).** Toda enzima que hidrolisa o RNA.
- Riborregulador (*riboswitch*).** Molécula de mRNA capaz de regular a expressão gênica – transcrição ou tradução – sofrendo alteração da conformação ao se ligar a um metabólito específico.
- Riborregulador (*riboswitch*) de lisina.** Um mRNA de bactérias que passa de uma conformação ativa (transcrita) a uma estrutura inativa (não transcrita) quando se liga à lisina.
- Ribossomo.** Organela citoplasmática na qual são sintetizadas as proteínas.
- RNA.** Ácido ribonucleico; material que contém informações em alguns vírus; de modo geral, uma molécula derivada do DNA por transcrição que pode levar informações (RNA mensageiro ou mRNA), garantir a estrutura subcelular (RNA ribossômico ou rRNA), transportar aminoácidos (RNA de transferência ou tRNA) ou facilitar a modificação bioquímica própria ou de outras moléculas de RNA.
- RNA antisense.** RNA complementar ao pré-mRNA ou mRNA produzido a partir de um gene.
- RNA de interferência curto (siRNA).** Moléculas de RNA bifilamentares com 21 a 28 pares de bases de comprimento que medeiam o fenômeno de interferência por RNA; também são conhecidas como moléculas de microRNA.
- RNA de transferência (tRNA).** RNA que transportam aminoácidos para os ribossomos, onde são usados para produzir proteínas.
- RNA-guia.** Molécula de RNA que contém sequências usadas como molde durante a edição do RNA.
- RNA mensageiro (mRNA).** RNA que leva do DNA para os ribossomos as informações necessárias para a síntese proteica.
- RNA nuclear heterogêneo (hnRNA).** A população de transcritos primários no núcleo de uma célula eucariótica.
- RNA polimerase.** Enzima que catalisa a síntese de RNA.



**RNA ribossômicos (rRNA).** Moléculas de RNA que são componentes estruturais dos ribossomos.

**RNA *sense*.** Transcrito primário de mRNA que contém uma região codificante (sequência contígua de códons), traduzida para produzir um polipeptídeo.

**Roentgen (r).** Unidade de radiação ionizante.

## S

**Saco embrionário.** Espaço grande e de paredes finas dentro do óvulo das espermatófitas no qual se desenvolvem a oosfera e, após fertilização, o embrião; o gametófito feminino maduro em plantas superiores.

**Salto no cromossomo (*chromosome jumping*).** Procedimento que usa grandes fragmentos de DNA para o deslocamento descontínuo de um sítio para outro ao longo de um cromossomo. (Veja também **Clonagem posicional**.)

**Segregação.** Separação dos cromossomos paternos e maternos na meiose; separação dos alelos em heterozigotos; ocorrência de diferentes fenótipos na prole, resultante da separação de cromossomos ou alelos nos pais heterozigotos; primeira lei de Mendel da hereditariedade.

**Seleção.** Sobrevivência e reprodução diferenciais entre os genótipos; os fatores mais importantes que modificam as frequências dos alelos em grandes populações.

**Seleção artificial.** A prática de escolher indivíduos de uma população para reprodução, geralmente porque esses indivíduos têm uma ou mais características desejáveis.

**Seleção em massa.** No cruzamento de vegetais e animais, a escolha de indivíduos para reprodução em toda a população tomando como base os fenótipos individuais, e não os fenótipos de seus parentes.

**Seleção genética.** Exposição de uma célula ou de um organismo a condições ambientais nas quais só é capaz de sobreviver se tiver um gene ou elemento genético específico.

**Seleção natural.** Sobrevivência e reprodução diferencial na natureza que favorecem indivíduos mais adaptados ao ambiente; eliminação de organismos menos aptos.

**Selenocisteína.** Aminoácido que contém selênio (número atômico 34) em vez de enxofre na cisteína.

**Selenoproteína.** Proteína que contém o aminoácido selenocisteína.

**Semidominante.** Termo designativo de alelos nos quais o fenótipo de um heterozigoto é intermediário entre os fenótipos dos homozigotos correspondentes.

**Semiesterilidade.** Condição de fertilidade apenas parcial em zigotos de vegetais (p. ex., milho); geralmente associada a translocações.

**Sequência -10.** Veja **Sequência TATAAT**.

**Sequência -35.** Veja **Sequência de reconhecimento**.

**Sequência CAAT (CAAT box).** Sequência nucleotídica conservada em promotores eucarióticos participantes do início da transcrição.

**Sequência da região 3' (*downstream*).** Sequência em uma unidade de transcrição depois do sítio de início da transcrição (mais próxima da extremidade 3'). O par de nucleotídeos no DNA correspondente ao nucleotídeo na extremidade 5' do transcrito (RNA) é designado +1. O par de nucleotídeos subsequente é designado +2. Todas as sequências de nucleotídeos subsequentes (+) são sequências da região 3' (cf. **Sequência da região 5'**).

**Sequência da região 5' (*upstream*).** Sequência em uma unidade de transcrição que precede o sítio de início da transcrição (está mais próxima da extremidade 5'). O par de nucleotídeos no DNA correspondente ao nucleotídeo na extremidade 5' do transcrito (RNA) é designado +1. O par de nucleotídeos precedente é designado -1. Todas as sequências de nucleotídeos precedentes (-) são sequências da região 5' (cf. **Sequência da região 3' [*downstream*]**).

**Sequência de ação *cis*.** Sequência de nucleotídeos que afeta apenas a expressão de genes localizados no mesmo cromossomo, isto é, *cis* em relação a si mesma.

**Sequência de consenso.** Sequência nucleotídica presente na maioria dos sinais ou elementos genéticos que desempenham uma função específica.

**Sequência de inserção.** Veja **Elemento IS**.

**Sequência de reconhecimento (sequência -35).** Sequência nucleotídica (consenso TTGACA) em promotores procarióticos à qual se liga o fator sigma da RNA polimerase durante o início da transcrição.

**Sequência de Shine-Dalgarno.** Sequência conservativa em mRNA procarióticos que é complementar a uma sequência próxima da terminação 5' do RNA ribossômico 16S e participa do início da tradução.

**Sequência líder.** Segmento de uma molécula de mRNA que vai da terminação 5' até o códon de iniciação da tradução.

**Sequência TATA (TATA box).** Sequência promotora conservada que determina o sítio de início da transcrição.

**Sequência TATAAT (sequência -10).** Sequência rica em AT nos promotores procarióticos que facilita o desenrolamento localizado do DNA e o início da síntese de RNA.

**Sequenciamento aleatório (*shotgun*) de genoma inteiro.** Técnica de sequenciamento de genomas que inclui a clivagem aleatória de todo o genoma em pequenos fragmentos, sequenciamento das extremidades desses fragmentos e uso de supercomputadores para montar a sequência completa por alinhamento de sequências superpostas.

**Sexo heterogamético.** Que produz gametas diferentes no que diz respeito aos cromossomos sexuais. Em seres humanos, o homem XY é heterogamético e a mulher XX é homogamética.

**Sexo homogamético.** Produz gametas iguais no que diz respeito aos cromossomos sexuais (cf. **Sexo heterogamético**).

**Sexodução.** Incorporação de genes bacterianos a fatores F e transferência subsequente, por conjugação, para uma célula receptora.

**Shelterina.** Complexo proteico que se liga aos telômeros e protege o DNA contra degradação.

**Silenciador.** Sequência de DNA que ajuda a reduzir ou bloquear a expressão de um gene adjacente.

**Simbionte.** Organismo que vive em íntima associação com outro organismo diferente.

**Sinal de término.** Na transcrição, uma sequência nucleotídica que especifica o término da cadeia de RNA.

**Sinapse.** Pareamento de cromossomos homólogos na prófase meiótica.

**Síndrome.** Grupo de sinais e sintomas que ocorrem juntos e representam uma doença específica.

**Síndrome de *cri-du-chat*.** Distúrbio provocado pela deleção de uma pequena região do braço curto de um cromossomo 5 humano.



- Síndrome de imunodeficiência adquirida.** Veja AIDS.
- Síndrome de imunodeficiência combinada grave (SCID).** Grupo de doenças caracterizadas pela incapacidade de resposta imune, seja humoral, seja celular.
- Síndrome de Klinefelter.** Distúrbio causado pela presença de dois cromossomos X e um cromossomo Y no cariótipo humano.
- SINE (elementos nucleares intercalados curtos).** Famílias de elementos transponíveis curtos (150 a 300 pb), moderadamente repetitivos, de eucariotos. A família SINE mais conhecida é a família Alu em seres humanos.
- Sintenia.** Ocorrência de dois *loci* no mesmo cromossomo, sem levar em conta a distância entre eles.
- Sítio aminoacil (A).** O sítio de ligação do ribossomo que recebe o influxo de aminoacil-tRNA.
- Sítio de restrição.** Sequência de DNA clivada por uma enzima de restrição.
- Sítio de saída (E).** O sítio de ligação do ribossomo que contém o tRNA livre antes de sua liberação.
- Sítio peptídil (P).** O sítio de ligação do ribossomo que contém o tRNA ao qual está ligada a cadeia polipeptídica em formação.
- Sítios hipersensíveis.** Regiões no DNA altamente suscetíveis à digestão por endonucleases.
- Sobredominância.** Condição na qual os heterozigotos são superiores (em alguma escala de medida) a qualquer um dos homozigotos associados a ele.
- Southern blot.** Transferência de fragmentos de DNA de gel de eletroforese para membrana de celulose ou náilon por ação capilar.
- SRY (região do Y determinante do sexo).** Gene ligado ao Y em seres humanos e outros mamíferos que codifica uma proteína, o fator determinante testicular, que tem papel fundamental no desenvolvimento masculino.
- STS (sítios marcados por sequência).** Sequências de DNA isoladas curtas (geralmente de 200 a 500 pb) que são amplificadas por PCR e usadas para ligar mapas físicos e mapas genéticos.
- Subespécie.** Uma entre duas ou mais populações diferentes do ponto de vista morfológico ou geográfico, mas intercruzadas de uma espécie.
- Substituição de base.** Substituição de uma base em uma molécula de DNA. (Veja também **Transição**; **Transversão**.)
- Substituição gênica.** Incorporação de um transgene em um cromossomo em sua localização normal por recombinação homóloga, assim substituindo a cópia do gene originalmente presente no *locus*.
- Substituição não sinônima.** Troca de um par de bases em um códon que altera o aminoácido especificado por esse códon.
- Substituição sinônima.** Substituição de um par de bases em um códon sem alteração do aminoácido especificado pelo códon.
- Super-hélice.** Molécula de DNA que contém torções adicionais em razão do superenrolamento (super-hélices positivas) ou subenrolamento (super-hélices negativas).
- Super-helicoidização negativa.** Formação de estruturas terciárias helicoidais em moléculas de DNA bifilamentares com extremidades fixas (que não são livres para girar) quando há subenrolamento das moléculas.

## T

- Taq polimerase.** DNA polimerase termoestável isolada da bactéria termofílica *Thermus aquaticus*.
- Taxa de concordância.** Entre os pares de itens identificados porque um membro do par tem determinada característica, a frequência com que o outro membro do par tem a mesma característica.
- Tay-Sachs, doença de.** Distúrbio autossômico recessivo, letal em seres humanos, caracterizado por degeneração neurológica e morte na fase inicial da infância. A doença é causada pela ausência da enzima hexosaminidase A.
- T-DNA.** Segmento de DNA no plasmídeo Ti de *Agrobacterium tumefaciens* que é transferido para células vegetais e inserido no cromossomo do vegetal.
- Telófase.** Último estágio de cada divisão mitótica ou meiótica no qual os cromossomos são montados nos polos do fuso de divisão.
- Telófase I.** Estágio da primeira divisão meiótica em que os cromossomos duplicados se reúnem no polo de uma célula em divisão e começam a se descondensar.
- Telófase II.** Estágio da segunda divisão meiótica em que os cromossomos se reúnem no polo de uma célula em divisão e começam a se descondensar.
- Telomerase.** Enzima que acrescenta sequências de telômeros às extremidades de cromossomos eucariotos.
- Telômero.** Estrutura especial encontrada na extremidade de cromossomos eucarióticos.
- Teoria cromossômica da hereditariedade.** Teoria que afirma que os cromossomos contêm as informações genéticas e que seu comportamento durante a meiose constitui a base física para a segregação e a distribuição independente de genes.
- Teoria neutra.** Teoria segundo a qual a evolução de características com efeito mínimo ou nulo sobre a aptidão é um processo aleatório que inclui mutação e deriva genética.
- Terapia gênica.** O tratamento de doenças hereditárias mediante introdução de cópias selvagens do gene defeituoso causador do distúrbio nas células de indivíduos afetados. Se as células reprodutivas forem modificadas, o procedimento é denominado *terapia gênica de células germinativas* ou *hereditária*. Se as células modificadas não forem reprodutivas, o procedimento é denominado *terapia gênica de células somáticas* ou *não hereditária*.
- Terapia gênica de células germinativas (hereditária).** Tratamento de um distúrbio hereditário pelo acréscimo de cópias ativas (selvagens) de um gene às células reprodutivas (germinativas) de um indivíduo que tem cópias defeituosas do gene em questão (cf. **Terapia gênica de células somáticas [não hereditária]**).
- Terapia gênica de células somáticas (não hereditária).** Tratamento de um distúrbio hereditário pelo acréscimo de cópias ativas (selvagens) de um gene às células não germinativas de um indivíduo que tem cópias defeituosas do gene em questão (cf. **Terapia gênica de células germinativas [hereditária]**).
- Terminalização.** Movimento de repulsão dos centrômeros de bivalentes nos estágios diplótenos da prófase meiótica que tende a mover os quiasmas visíveis em direção às extremidades dos bivalentes.



- Término (da síntese de DNA, RNA ou proteínas).** Liberação de uma macromolécula completa (DNA, RNA ou polipeptídeo) após a incorporação da última subunidade (nucleotídeo ou aminoácido).
- Teste *cis-trans*.** Construção e análise de heterozigotos *cis* e *trans* de pares de mutações para verificar se as mutações estão no mesmo gene ou em dois genes diferentes. Para que o teste seja informativo, é preciso que o heterozigoto *cis* tenha o fenótipo selvagem. Se essa condição for atendida, as duas mutações estão no mesmo gene se o heterozigoto *trans* tiver o fenótipo mutante e estão em dois genes diferentes se o heterozigoto *trans* tiver o fenótipo selvagem.
- Teste de complementação (teste *trans*).** Introdução de duas mutações recessivas na mesma célula para determinar se são alelos do mesmo gene, ou seja, se afetam a mesma função genética. Se as mutações forem alélicas, o genótipo  $m_1 +/+ m_2$  exibirá um fenótipo mutante; caso contrário, o fenótipo será selvagem.
- Testosterona.** Hormônio esteroide que induz o desenvolvimento de características masculinas.
- Tétrade.** As quatro células originadas da segunda divisão meiótica em vegetais (tétrades de pólen) ou fungos (ascósporos). O termo também é usado para identificar o grupo de quatro cromátides formado pela associação de cromossomos homólogos duplicados durante a meiose.
- Tetraploide.** Organismo cujas células contêm quatro conjuntos haploides ( $4n$ ) de cromossomos ou genomas.
- Tetrassômico.** Relativo a um núcleo ou organismo que tem quatro membros de um de seus cromossomos, enquanto o restante do complemento cromossômico é diploide. (Fórmula cromossômica:  $2n + 2$ .)
- TFIIIX (fator de transcrição X para RNA polimerase II).** Proteína necessária para o início da transcrição por RNA polimerase II em eucariotos; X representa qualquer um dos diferentes fatores designados A a F.
- Timina (T).** Base pirimidínica encontrada no DNA. As três outras bases orgânicas – adenina, citosina e guanina – são encontradas tanto no RNA quanto no DNA, mas no RNA, a timina é substituída por uracila.
- Tipo selvagem.** O fenótipo habitual ou padrão para comparação.
- Topoisomerase.** Enzima que introduz ou remove super-hélices do DNA.
- Tradução.** Síntese de proteína (polipeptídeo) direcionada por um RNA mensageiro específico; ocorre em ribossomos.
- Tráfego.** Movimento de substâncias através do citoplasma de uma célula, geralmente guiado por membranas, vesículas e componentes do citoesqueleto.
- Transatante.** Termo que descreve substâncias difusíveis e capazes de afetar estruturas celulares separadas no espaço.
- Transcrição.** Processo de formação de RNA ao longo de um molde de DNA. A enzima RNA polimerase catalisa a formação de RNA a partir de trifosfatos de ribonucleosídeo.
- Transcriptase reversa.** Enzima que catalisa a síntese de DNA usando um molde de RNA.
- Transcriptoma.** O conjunto completo de RNA transcritos de um genoma.
- Transcrito.** Molécula de RNA produzida por transcrição de um gene.
- Transcrito primário.** Molécula de RNA produzida por transcrição antes de qualquer modificação pós-transcricional; também denominada pré-mRNA em eucariotos.
- Transdução (t).** Recombinação genética em bactérias mediada por bacteriófago. Transdução abortiva: o DNA bacteriano é injetado por um fago em uma bactéria, mas não se replica. Transdução generalizada: qualquer gene bacteriano pode ser transferido por um fago para uma bactéria receptora. Transdução restrita: a transferência de DNA bacteriano por um fago temperado é restrita a apenas um sítio no cromossomo bacteriano.
- Transdução de sinal.** Processo pelo qual um sinal molecular, como um hormônio, passa internamente dentro de uma célula por um sistema de moléculas para alterar o estado celular.
- Transdução especializada.** Recombinação em bactérias mediada por um bacteriófago que só é capaz de transferir genes em um pequeno segmento do cromossomo da célula doadora para uma célula receptora (cf. **Transdução generalizada**).
- Transdução generalizada.** Recombinação em bactérias mediada por um bacteriófago capaz de transferir qualquer gene bacteriano da célula doadora para uma célula receptora (cf. **Transdução especializada**).
- Transfecção.** Captação de DNA por uma célula eucariótica, seguida pela incorporação ao genoma celular de marcadores genéticos presentes no DNA.
- Transferase terminal.** Enzima que acrescenta nucleotídeos às terminações 3' das moléculas de DNA.
- Transformação (cancerosa).** Conversão de células eucarióticas que crescem em cultura em um estado de crescimento celular descontrolado (semelhante ao crescimento de células tumorais).
- Transformação (genética).** Alteração genética de um organismo desencadeada pela incorporação de DNA estranho às células.
- Transformação mediada por *Agrobacterium tumefaciens*.** Processo natural de transferência de DNA da bactéria *A. tumefaciens* para vegetais.
- Transgene.** Gene estranho ou modificado que foi introduzido em um organismo.
- Transgênico.** Termo aplicado a organismos que foram modificados por introdução de moléculas de DNA.
- Transição alostérica.** Interação reversível de uma molécula de proteína com uma pequena molécula que modifica o formato da proteína e, conseqüentemente, altera a interação dessa proteína com uma terceira molécula.
- Transição.** Mutação causada pela substituição de uma purina por outra purina ou de uma pirimidina por outra pirimidina no DNA ou RNA.
- Translocação.** Troca de posição de um segmento de um cromossomo para outra parte do mesmo cromossomo ou para outro cromossomo.
- Translocação robertsoniana.** Rearranjo no qual os braços longos de dois cromossomos homólogos se uniram nos centrômeros, ou perto deles, e os braços curtos desses cromossomos se perderam.
- Transposase.** Enzima que catalisa o deslocamento de uma sequência de DNA para outro sítio em uma molécula de DNA.



- Transpóson composto.** Elemento transponível formado quando dois transpósons idênticos ou quase idênticos se inserem de cada lado de um segmento não transponível de DNA – por exemplo, o transpóson bacteriano Tn5.
- Transpóson de “cortar e colar”.** Elemento transponível que é excisado de uma posição no genoma e inserido em outra pela ação da transposase, uma enzima codificada por transpóson.
- Transpóson replicativo.** Elemento transponível replicado durante o processo de transposição. Tn3 em *E. coli* é um exemplo.
- Transpósons.** Elementos de DNA capazes de se deslocarem (“transpor”) de uma posição para outra em uma molécula de DNA.
- Transversão.** Mutação causada pela substituição de uma pirimidina por uma purina ou de uma purina por uma pirimidina no DNA ou RNA.
- 2', 3'-Trifosfatos de dideoxirribonucleosídeo (ddNTPs).** Precursores de DNA de término da cadeia (trifosfatos de nucleosídeos) com um átomo de hidrogênio (H) ligado ao carbono 3' em lugar do grupo hidroxila (OH) observado em precursores de DNA normais (2'-trifosfato de desoxirribonucleotídeo); os ddNTPs são usados em reações de sequenciamento do DNA.
- Tri-híbrido.** A prole de pais homozigotos que diferem em três pares de genes.
- Trissômico.** Célula ou organismo diploide que tem um cromossomo a mais em um par (fórmula cromossômica:  $2n + 1$ ). Um caso específico desse distúrbio é denominado trissomia.
- Trivalente.** Associação entre três cromossomos durante a meiose.
- tRNA supressor.** tRNA mutante que reconhece um ou mais dos códons de término e insere um aminoácido em um sítio onde normalmente a tradução terminaria.
- tRNA<sub>f</sub><sup>Met</sup>.** O tRNA de metionina que especifica a iniciação de cadeias polipeptídicas em procariotos (cf. tRNA<sub>f</sub><sup>Met</sup>, tRNA<sub>i</sub><sup>Met</sup>).
- tRNA<sub>i</sub><sup>Met</sup>.** O tRNA de metionina que especifica a iniciação de cadeias polipeptídicas em eucariotos (cf. tRNA<sub>f</sub><sup>Met</sup>, tRNA<sub>i</sub><sup>Met</sup>).
- tRNA<sup>Met</sup>.** O tRNA de metionina que responde aos códons internos de metionina, e não aos códons de iniciação (cf. tRNA<sub>f</sub><sup>Met</sup>, tRNA<sub>i</sub><sup>Met</sup>).
- Tubulina.** Principal proteína constituinte dos microtúbulos de células eucarióticas.
- Turner, síndrome de.** Fenótipo associado ao genótipo XO em seres humanos.

## U

- Unidade de crossing over.** Medida da distância em mapas genéticos baseada no número médio de eventos de *crossing over* ocorridos durante a meiose. Um intervalo de mapa com comprimento de uma unidade de *crossing over* (às vezes denominada centiMorgan) significa que só uma em cada cem cromátides recuperadas da meiose sofreram um evento de *crossing over* nesse intervalo.
- Unidade de mapa.** Veja Unidade de *crossing over*.
- Unidade de transcrição.** Segmento de DNA que contém sinais de início e término da transcrição e é transcrito em uma molécula de RNA.

**Univalente.** Cromossomo não pareado na meiose.

**Universalidade (do código genético).** Os códons têm o mesmo significado, com pequenas exceções, em praticamente todas as espécies.

**Uracila (U).** Base pirimidínica encontrada no RNA, mas não no DNA. No DNA, a uracila é substituída por timina.

## V

- Valor crítico.** Valor limiar de uma estatística que define uma fração da distribuição de frequência da estatística. Uma estatística por amostragem maior que esse valor crítico confirma a rejeição da hipótese testada.
- Valor médio dos pais.** Em genética quantitativa, a média dos fenótipos dos pais.
- van der Waals, interações de.** Atrações fracas entre átomos muito próximos.
- Varição.** Em biologia, a ocorrência de diferenças entre indivíduos.
- Varição contínua.** Variação não representada por classes distintas. Os indivíduos misturam-se e a análise requer dados de medida (cf. **Varição descontínua**). Em geral, múltiplos genes são responsáveis por esse tipo de variação.
- Varição descontínua.** Variabilidade fenotípica de classes distintas como a cor vermelha ou branca, estatura elevada ou anã (cf. **Varição contínua**).
- Varição transgressiva.** Aparecimento na geração F<sub>2</sub> (ou posterior) de indivíduos que apresentam desenvolvimento mais extremo de uma característica do que os pais.
- Variância genética aditiva.** A proporção da variância fenotípica total em uma característica quantitativa decorrente dos efeitos aditivos de alelos.
- Variância.** Medida da variação em uma população; o quadrado do desvio padrão.
- Variagem por efeito de posição.** Variação fenotípica em um indivíduo causada por alteração na posição genômica de um gene. Em geral, esse tipo de variação é observado quando um gene naturalmente localizado na eucromatina é deslocado por um rearranjo cromossômico para uma região heterocromática do genoma.
- Vetor bifuncional.** Plasmídeo capaz de se replicar em dois organismos diferentes, como leveduras e *E. coli*.
- Vetor de clonagem.** Molécula de DNA pequena, capaz de autorreplicação – geralmente um plasmídeo ou um cromossomo viral – na qual são inseridos DNA estranhos durante o processo de clonagem de genes ou de outras sequências de DNA de interesse.
- Vetor.** Plasmídeo ou cromossomo viral que pode ser usado para construir moléculas de DNA recombinantes a serem introduzidas em células vivas.
- Via anabólica.** Via de síntese de um metabólito; via de biossíntese.
- Via catabólica.** Via pela qual uma molécula orgânica é degradada para obter energia para o crescimento e os outros processos celulares; via de degradação.
- Viabilidade.** Capacidade de viver e se desenvolver normalmente.
- Vigor híbrido (heterose).** Crescimento, força e saúde incomuns de híbridos heterozigotos derivados de dois homozigotos menos vigorosos.



**Vírus da imunodeficiência humana (HIV).** Retrovírus causador da síndrome de imunodeficiência adquirida (AIDS) em seres humanos.

**VNTR.** Veja **Repetição em série em número variável**.

## W

**Western blot.** Transferência de proteínas de gel de eletroforese para membrana de celulose ou náilon por ação de uma força elétrica.

## Y

**YAC (cromossomos artificiais de levedura).** Vetores de clonagem lineares elaborados a partir de elementos essenciais de cromossomos de levedura. Conseguem acomodar inserções de DNA estranho de 200 a 500 kb.

## Z

**Z-DNA.** Dupla hélice levógira que se forma em moléculas de DNA ricas em GC. O Z refere-se a disposição em zigue-zague da cadeia principal de açúcar-fosfato nessa forma de DNA.

**Zigóteno (zigonema).** Estágio da meiose em que ocorre a sinapse; sucede o leptóteno e precede o paquíteno na prófase meiótica.

**Zigoto.** Célula produzida pela união de duas células sexuais maduras (gametas) na reprodução; também designa, em genética, o indivíduo que se desenvolve a partir dessa célula.